Causation in the Nature-Nurture Debate: The Case of Genotype-Environment Interaction

by

James Gregory Tabery

Bachelor of Science, Fairfield University, 2000

Submitted to the Graduate Faculty of

Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2007

UNIVERSITY OF PITTSBURGH

COLLEGE OF ARTS AND SCIENCES

This dissertation was presented

by

James Gregory Tabery

It was defended on

July 17, 2007

and approved by

Robert C. Olby, Research Professor, Dept. of History and Philosophy of Science

Michael Pogue-Geile, Associate Professor, Psychology Department

Kenneth F. Schaffner, University Professor, Dept. of History and Philosophy of Science

Dissertation Director: Sandra D. Mitchell, Professor, Dept. of History and Philosophy of Science

Dissertation Director: Paul E. Griffiths, Professor, Philosophy, University of Queensland

Causation in the Nature-Nurture Debate: The Case of Genotype-Environment Interaction

James Gregory Tabery

University of Pittsburgh, 2007

I attempt to resolve an aspect of the nature-nurture debate. Consider a typical nature-nurture question: Why do some individuals develop a complex trait such as depression, while others do not? This question incorporates an etiological query about the *causal mechanisms responsible for the individual development* of depression; it also incorporates an etiological query about the *causes of variation responsible for individual differences* in the occurrence of depression. Scientists in the developmental research tradition of biology investigate the former; scientists in the biometric research tradition of biology investigate the latter. So what is the relationship?

The developmental and biometric research traditions, I argue, are united in their joint effort to elucidate what I call *difference mechanisms*. Difference mechanisms are regular causal mechanisms made up of difference-making variables that take different values in the natural world. On this model, individual differences are the effect of difference-makers in development that take different values in the natural world.

I apply this model to the case of genotype-environment interaction (or G×E), showing that there have actually been two separate concepts of G×E: a *biometric* concept (or $G\times E_B$) and a *developmental* concept (or $G\times E_D$). These concepts also may be integrated via the difference mechanisms model: G×E results from the interdependence of difference-makers in development that take different values in the natural world.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

This dissertation is about difference-makers in development that take different values. There have been a great number of difference-makers in the development of this dissertation, and I would like to acknowledge their value. As co-directors of the dissertation, **Paul Griffiths** and **Sandra Mitchell** were the driving forces behind seeing the product from start to finish. I can still recall Professor Griffiths handing me a copy of Hogben's *Nature and Nurture* some four years ago; in many ways, the dissertation began with that transaction. And I can still hear Professor Mitchell encouraging me to insert my own voice into the analysis; her urging pushed the dissertation from a descriptive, historical analysis to a critical, philosophical analysis. I also benefited from very involved committee members. **Robert Olby** taught me how to do history of biology, and I can think of no one else who I would prefer having learned it. **Michael Pogue-Geile** graciously took an interest in an historical and philosophical analysis of his scientific discipline, and I hope I have made his effort worth the while. Finally, **Kenneth Schaffner** read multiple drafts of every chapter and provided me with invaluable advice all along the way; his personal interest in the nature-nurture debate was an affirmation for me of the project's importance.

Conversations with a number of historians, philosophers, and scientists also made a difference in the project's development: **André Ariew**, **William Bechtel**, **Zvi Biener**, **James Bogen**, **Mark Borrello**, **Ingo Brigandt**, **Avshalom Caspi**, **Roderick Cooper**, **Carl Craver**, **Thomas Cunningham**, **Lindley Darden**, **Megan Delehanty**, **Evelyn Fox Keller**, **Peter Gildenhuys**, **Gilbert Gottlieb**, **James Griesemer**, **Adolf Grünbaum**, **Balázs Gyenis**, **Andrew Hamilton**, **James Lennox**, **Alan Love**, **Peter Machamer**, **Edouard Machery**, **Roberta Millstein**, **Terrie Moffitt**, **David Moore**, **Lisa Parker**, **Kathryn Plaisance**, **Robert Plomin**, **Richard Purcell**, **Michael Rutter**, **Elliott Sober**, **Jacqueline Sullivan**, **C. Kenneth Waters**, and **Christian Wüthrich**. Archivists at the American Philosophical Society, the University of Adelaide, the University of Birmingham, and University College London either welcomed me into their libraries or made available to me items in their archive, and Leslie Hogben kindly permitted me to quote from the letters from her grandfather. Versions of the dissertation chapters also were presented at Pomona College (January 2007, Claremont, CA), Rice University (January 2007, Houston, TX), Na-Nu (March 2007, Indiana University, Bloomington, IN), the 3rd Queensland Biohumanities Project (December 2006, University of Queensland, Brisbane, Australia), the Southwest Conference for the History and Philosophy of the Life Sciences (March 2006, UC-Davis, Davis, CA), the History of Science Society's annual meeting (November 2005, Minneapolis, MN), the British Society for the History of Science's annual meeting (July 2005, Leeds, UK), the International Society for the History, Philosophy, and Social Studies of Biology's biannual meeting (July 2005, Guelph, Ontario), the Canadian Society for the History and Philosophy of Science's annual meeting (May 2005, London, Ontario), and *Beyond Dichotomies, Across Boundaries* (April 2005, University of Minnesota, Minneapolis, MN). Conversations with a number of conference participants helped me to clarify ideas on the topic.

Fortunately, my time in Pittsburgh was not entirely spent writing the dissertation. Members of the History and Philosophy of Science Department, the Center for Bioethics and Health Law, and the Center for Philosophy of Science have made Pittsburgh a wonderful place to study, learn, work, and live. In particular, dear roommates including **Christian Wüthrich**, **David Miller**, and **Brian Hepburn**, and dear friends including **Thomas and Stephanie**

Cunningham and **Richard Purcell** *almost* make me wish the dissertation was not yet complete…almost.

I owe the greatest debt of gratitude to my parents: **Michael Tabery** and **Pamela Tabery**, as well as my brother **Jeff Tabery**. My parents provided me with my genotype and, more importantly, the supportive, nurturing environment with which my genotype could interact. The dissertation is dedicated to them.

PREFACE

This dissertation is an exercise in *integration*. It is integrative in multiple senses—both in terms of the *product* and in terms of the *tools* employed to create that product. On the one hand, the product is an integration. In the early, historical chapters of the dissertation (Chapters 2 and 3), I trace out two separate research traditions that both attempt to elucidate the etiology of complex traits: a *biometric research tradition* and a *developmental research tradition*. These two traditions have largely been at odds during the history of the nature-nurture debate, competing over how to best address the problem of elucidating the etiology of complex human traits. Examining the nature-nurture debate in its entirety would be much too large of a project for any single dissertation, but it is possible to trace a thread within that larger debate. The thread that I trace is disputes over genotype-environment interaction, or G×E. I show that, just as there have been two separate research traditions involved in the debates, so too have there been two separate concepts of G×E, each with its own legacy: what I call the *biometric* concept, or $G\times E_B$, and what I call the *developmental* concept, or $G\times E_D$. The purpose of the historical chapters is to tease apart the epistemological and methodological components of the separate research traditions, and show how the separate concepts of G×E are situated in those separate traditions. In the later, philosophical chapters (Chapters 4 and 5), I bring these separate research traditions and these separate concepts back together. I integrate them. The biometric and developmental research traditions are integrated based upon what I identify as a shared problem of interest—the elucidation of what I call *difference mechanisms*. And the biometric and developmental concepts of G×E are integrated based upon a concept related to difference mechanisms—*the interdependence of difference-makers in development that take different values in the natural world*. It is in this sense that the product of this dissertation is an integration.

The tools employed to create that product also are integrated. This is a work in history AND philosophy of science, or what some are now calling *integrative* history and philosophy of science—&HPS.[1] The dissertation can be roughly broken down into two historical chapters and two philosophical chapters, but this division belies the fact that I employ both historical and philosophical tools throughout the dissertation. Resolving the debates over G×E (or the nature-nurture debate more broadly) demands this two-pronged approach. These debates have been fought for nearly a century now; the sides were partitioned from the start, and they persist into the present. The tools of the historian are required to track this debate, to situate the various disputants in their respective disputes, noting how the debates have changed or stayed the same over time. Importantly, though, the debates are fundamentally over philosophical issues: causation and explanation. So it is not enough to just situate the disputants in their respective disputes. If I am to engage these debates myself, then the tools of the philosopher are also essential to evaluate the claims being made on each side. I attempt to track the debates with the tools of the historian (e.g., archival research, interviews with retired and practicing scientists, critical engagement with primary sources). I attempt to resolve the debates with the tools of the philosopher (e.g., a disentanglement of the epistemological and methodological components of scientific research traditions, a conceptual distinction, an analysis of causal-mechanical explanation). It is in this sense that the tools of this dissertation are integrated.

[1] &HPS is the brain-child of John D. Norton and Don Howard. The organization currently consists of these two conveners as well as nearly twenty committee members, all of whom are highly respected members of the history of science and the philosophy of science communities. The first conference devoted to &HPS—&HPS1—will be held at the University of Pittsburgh, October 11-13, 2007.

1. INTRODUCTION

Despite the widely endorsed "interactionist credo" (Kitcher 2001, 398), the nature-nurture debate remains a quagmire of epistemological and methodological disputes over causation, explanation, and the concepts employed therein. The nature-nurture debate at its most fundamental level is a debate about the etiology of phenotypic traits. It is a debate about how science answers questions such as the following: Why do some individuals have a better memory than other individuals? Or, why do some individuals develop a complex trait such as depression, antisocial behavior, or schizophrenia, while others do not?

Notice that these questions incorporate an etiological question about the *regular causal mechanisms responsible for the individual development* of the trait, and they also incorporate an etiological question about the *causes of variation responsible for individual differences* in the trait. This distinction between the regular causal mechanisms responsible for individual development and the causes of variation responsible for individual differences has been at the heart of the disputes over how to answer these etiological questions.

Etiological questions about complex behavioral traits are asked by a number of different scientists; developmental biologists, molecular biologists, neurobiologists, population geneticists, psychiatrists, psychologists, quantitative behavioral geneticists, sociologists, to name just a few, all contribute to the enterprise. For a century now, though, the persistent tension has been between biometrically-oriented scientists and developmentally-oriented scientists. Scientists in the biometric research tradition, such as early population geneticists or contemporary quantitative behavioral geneticists, seek to elucidate the causes of variation responsible for individual differences in a population; they ask *how-much?* questions about the causes of variation that are responsible for these individual differences; and they utilize statistical

methodologies such as the analysis of variance to answer the questions. Scientists in the developmental research tradition, such as early experimental embryologists or contemporary developmental psychobiologists, seek to elucidate the regular causal mechanisms responsible for individual development; they ask *how?* questions about this causal process; and they utilize interventionist methodologies to answer the questions (Table 1).[2]

Components	**Biometric Tradition**	**Developmental Tradition**
Problem	Individual Differences	Individual Development
Approach to Causation	Causes of Variation	Causal Mechanisms
Causal Question	How Much?	How?
Methodology	Statistical	Interventionist

Table 1. The components of the biometric and developmental research traditions.

In converging on the common domain of etiological questions about complex traits, then, disputes have arisen between scientists in the biometric tradition and scientists in the developmental tradition over how to best answer such questions. Often these disputes have revolved around *genotype-environment interaction*, or G×E.

1.1. Genotype-Environment Interaction

The concept of G×E refers to cases where different genotypic groups phenotypically respond differently to the same array of environments. Such interactions are often described in terms of genetic *predisposition*, genetic *risk*, genetic *liability*, or genetic *propensity* to the particular trait under investigation, and emphasize the fact that what one genotype does in an array of

[2] I will explicate the separate components of the biometric and developmental research traditions as well as explore the relationship between the separate components in the subsequent chapters.

environments can be quite different from what another genotype does in that same array of environments (Plomin et al. 1997, 269).[3] The standard method for representing a genotype's phenotypic response to a varying environment is by way of a reaction norm (or norm of reaction) graph.[4] For example, Figure 1 depicts two reaction norms from a recent study, one for individuals with genotypes conferring low MAOA activity and one for individuals with genotypes conferring high MAOA activity exposed to varying degrees of childhood maltreatment (None, Probable, and Severe) and leading to differences in risk of developing antisocial behavior (Caspi et al., 2002). MAOA (monoamine oxidase A) is a neurotransmitter-metabolizing enzyme, which inactivates neurotransmitters such as dopamine, norepinephrine, and serotonin (Shih, Chen, and Ridd 1999). Deficiencies in this enzyme have been linked with aggression (Rowe 2001). While individuals from both groups tend to increase in risk of developing antisocial behavior in response to increased childhood maltreatment, those individuals with genotypes conferring low MAOA activity actually score *lower* than individuals with genotypes conferring high MAOA activity on an antisocial behavior index when there is no reported childhood maltreatment, score *slightly higher* when there is probable childhood maltreatment, and score *much higher* when there is severe childhood maltreatment. This variable phenotypic response of different genotypes to the varying environments is the hallmark of G×E.

[3] For certain cases of G×E, concepts such as *genetic* predisposition and *genetic* propensity fundamentally misconstrue the phenomenon. I argue elsewhere that a new concept—*interactive predisposition*—must be introduced to capture these cases of G×E as well as the unique ethical implications raised by these cases (Tabery 2007).

[4] For a history of the reaction norm concept, see Sarkar (1999).

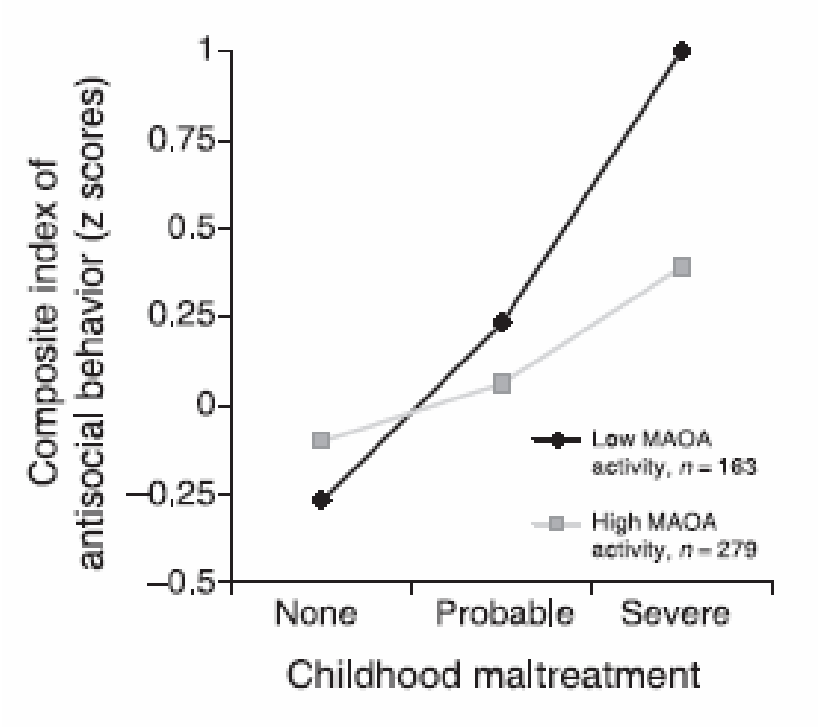

Figure 1. Norms of reaction for genotypes conferring low and high MAOA activity in response to varying levels of childhood maltreatment. From Caspi et al. (2002, Figure 1).

Cases of G×E have important implications for the study of the etiology of traits. First, if G×E exists for a particular trait in a population, then a scientist cannot assume that phenotypic variation for that trait in a population is simply the sum of genotypic differences and environmental differences (the "main effects"). The presence of G×E adds another source of variation which must be taken into consideration. If no G×E exists, then an "additivity relation" may be assumed, and the statistical analysis of variance (ANOVA) may be employed to partition the total phenotypic variance (V_P) into genotypic variance (V_G) and environmental variance (V_E):[5]

[5] For the additivity relation to hold, there also must be no genotype-environment correlation (rGE). In contrast to G×E, rGE refers to cases where there is a correlation between a genotype's *exposure* to particular environments. This form of interdependence between genotype and environment also generates its own source of variation.

$$V_P = V_G + V_E \quad (1.1)$$

When additivity applies, we can talk also about the *proportion* of total phenotypic variation attributable to either genotypic variation or environmental variation. For example, the concept of broad heritability (h^2), often described as the "estimation of the degree of genetic determination" (Falconer and Mackay 1996, 123), is measured as:

$$h^2 = V_G/V_P \quad (1.2)$$

But when the effect of genetic differences is modified by the environmental distribution and the effect of environmental differences is modified by genetic distribution, Equation (1.1) must be modified to include variation due to G×E ($V_{G\times E}$):

$$V_P = V_G + V_E + V_{G\times E} \quad (1.3)$$

Now a heritability measure can not be estimated because the genotypic and environmental variances are no longer independent. It may be possible to get around this complication by changing the scale on which the variables are measured in order to get back to an additivity relation (a transformation of scale). However, this transformation of scale requires a justification in itself. If it is employed purely for the sake of statistical convenience without regard to any plausible biological framework, then it is unclear what biological information the measurement provides after the transformation has been performed (Hernandez and Blazer 2006, chapter 8). Moreover, a transformation of scale may not always even be possible. Cases of G×E result in either a *change of rank* or a *change of scale* (Lynch and Walsh 1997, 658). In cases of a change of rank, the reaction norms actually cross, as is the case in Figure 1. In cases of a change of scale, the higher ranking genotype in one environment only reacts more or less strongly to the conditions of other environments but maintains its higher ranking in the other environments (for

an example, see Figure 7 in Chapter 2). A transformation of scale is not possible for the cases of G×E resulting in a change of rank (Lynch and Walsh 1997, 679).

Second, and related to the first point, since instances of G×E can be to such a degree that norms of reaction actually change rank across different environments, then it becomes clear that even though one genotypic group may perform better than another genotypic group in one environment does not necessarily mean that this will be the case in other environments. As a result, scientists must be wary of inferences made about the performance of different genotypic groups in untested environments simply from the knowledge of how those groups performed in limited, tested environments (Falconer and Mackay 1996, 132-134).[6]

1.2. Genotype-Environment Interaction in the Nature-Nurture Debate

The nature-nurture debate has been many things over many years. At its heart, though, the debate is over what role genes (nature) and the environment (nurture) play in the etiology of complex human traits. If the causal mechanisms responsible for complex human traits were elucidated, then much of the debate would be resolved.[7] We could point to the description of the elucidated mechanisms and show what role genes play, what role the environment plays, and how these two variables interact in the development of the traits. But such causal-mechanical explanations are far from complete. For ethical reasons, scientists cannot easily manipulate genetic and environmental variables in humans by, say, selectively breeding humans or drastically altering their environments with deprivation studies.

[6] These two "implications" paragraphs are drawn from Falconer and Mackay's explication of G×E in their textbook, *Introduction to Quantitative Genetics* (1996). Such discussions are now standard in any quantitative genetics textbook (see also Kearsey and Pooni 1996, chapter 12; Lynch and Walsh 1997, chapter 22).

[7] Well, at least the scientific debate would be resolved. There would still be reason to debate how a society should respond to this scientific knowledge, and this societal response is arguably also part of the nature-nurture debate.

When such causal-mechanical explanations are not available, what are scientists interested in the etiology of complex human traits to do? For a century, scientists in the biometric research tradition have turned to populations and attempted to determine how much of the total variation in a trait like depression is attributable to different causes of variation with statistical tools such as ANOVA. The results, through much of the twentieth century, were statistical heritability estimates (see equation 1.2 above) that apportioned the proportion of variation in the population attributable to genetic differences and utilized that measure to reflect, as mentioned above, "the degree of genetic determination."

But, as was already mentioned, if genetic and environmental differences are interdependent, then there is G×E. And with G×E, a serious complication arises for the statistical estimates. G×E, as a result, has been situated at the heart of the nature-nurture debate. Noting the interactive relationship between genes and the environment during individual development, critics of the biometric research tradition have attacked their statistical tools, which attempt to partition separate causes of variation with a focus on main effects. Such separation of genetic and environmental causes, the critics argue, misses the interdependence between these causes during the developmental process. But the scientists in the biometric tradition, in reply, contend that the developmentally-oriented scientists misunderstand G×E. G×E, they argue, is a statistical concept pertaining to causes of variation responsible for individual differences and having nothing to do with individual development. When G×E does exist, they conclude, their statistical tools can detect it, measure it, and often even eliminate the complication with a transformation of scale, so the attacks from a developmentally-conceived G×E misunderstand both the concept and the methodology of the biometric research tradition.

The debate over G×E is most obviously a dispute over what this concept actually captures in the natural world. Is it simply a population-level, statistical measure—a breakdown in the additivity of main effects? Or does it incorporate something about the developmental interactions between genes and the environment? Importantly, however, the debate is more than just about *defining* the concept. Those who identify a relationship between G×E and developmental interactions argue that such instances of G×E should be sought out for the light they shed on development and variation; they criticize the scientists in the biometric tradition for either ignoring the problem or attempting to eliminate the nuisance with a transformation of scale. When scientists in the biometric tradition find no evidence of G×E with ANOVA, the critics point out that this is just an example of how poorly the statistical methodologies inform us about the complexities of the natural world. But those who conceive of G×E as simply a statistical measure reply that the concept has nothing to do with the complexities of individual development. They argue that their statistical methodology works just fine for examining individual differences in populations, and they conclude that when ANOVA detects no G×E, this is a reflection of the lack of interaction in the natural world, not a reflection of ANOVA's mishandling of the phenomenon.

In sum, the debate over G×E is just one facet of the complicated and perennial nature-nurture debate. That said, it has resided at the heart of the epistemological, methodological, and conceptual disputes between the scientists in the biometric research tradition who, recognizing the present inability to elucidate the causal mechanisms of development, have turned to statistical data to answer the etiological questions about complex human traits, and the scientists in the developmental research tradition who have attacked the simplifications of genetic and environmental causation that have apparently followed from this statistical turn. The job for the

historian and philosopher of science, then, is to sort out these epistemological, methodological, and conceptual disputes with the tools of the trade.

I attempt this sorting-out in two basic steps: a disentangling, and a re-integration. In Chapters 2 and 3 I disentangle the various axes upon which the nature-nurture debate has been disputed. These are the various components of the biometric and the developmental research traditions: *problems* of focus, *approaches to causation* designed to resolve those problems, *causal questions* asked about those problems, and *methodologies* employed to provide the answers. I also disentangle the different concepts of G×E in these chapters, showing how the concepts emerged in and persisted through the separate research traditions. With the separate research traditions and the separate concepts disentangled, I then turn in Chapters 4 and 5 to re-integrating these elements by explicating the actual relationships between the traditions and the concepts. The biometric and developmental research traditions are integrated via what I call *difference mechanisms* in Chapter 4. And the biometric and developmental concepts of G×E are integrated in Chapter 5 via a concept related to difference mechanisms—*the interdependence of difference-makers in development that take different values in the natural world*.

1.3. $G \times E_B$ vs. $G \times E_D$

The first half of the dissertation is largely historical in nature. I trace the debates over G×E from their origin*(s)* to the present. I say "origin*(s)*" and not "the origin" because I show in Chapter 2 that there have actually been two distinct concepts of G×E since the earliest research on the phenomenon: a *biometric* concept, or $G \times E_B$, and a developmental concept, or $G \times E_D$. The biometric concept was introduced by R. A. Fisher, one of the fathers of population genetics and the creator of ANOVA. Fisher, operating in the biometric research tradition, was attempting to

develop statistical methods to measure the relative contributions of nature and nurture in the 1920's and 1930's. Fisher's consideration of G×E grew out of this biometric research program, as he increasingly realized the problems posed by G×E for his partitioning of causes of variation. $G\times E_B$, for Fisher, was simply a statistical measure of the breakdown in additivity between genotype and environment to be eliminated with a transformation of scale (if it even existed at all). Fisher, however, was not the only British biologist and statistician considering G×E in the beginning of the twentieth century. Lancelot Hogben introduced the developmental concept of G×E as he was considering the role that development played in the generation of variation in a population. Hogben was operating in the developmental research tradition. $G\times E_D$, for Hogben, was a result of differences in unique developmental combinations of genotype and environment. After tracking Fisher and Hogben's separate routes to G×E, I draw on this history to sort out the debate between Fisher and Hogben over the importance of such interactions and also trace $G\times E_B$ and $G\times E_D$ beyond Fisher and Hogben into mid-twentieth century population and developmental genetics.

Chapter 3 is devoted to exploring the place of G×E in the IQ Controversy, probably the debate over G×E most familiar to historians and philosophers of science. The IQ Controversy began in 1969 with the publication of Arthur Jensen's "How Much Can We Boost IQ and Scholastic Achievement" (Jensen 1969), where Jensen drew on heritability estimates for IQ to attribute the gap in IQ scores between black and white populations to genetic differences. Following in the footsteps of Hogben and developmental geneticists such as Conrad Hal Waddington, developmentally-oriented scientists such as Richard Lewontin and David Layzer attacked Jensen's heritability estimates with arguments based on a developmentally-conceived G×E, while Jensen deflected such attacks by arguing that G×E had nothing to do with

development and replying that ANOVA can detect G×E if and when it exists and eliminate it with a transformation of scale. After delineating these separate positions, I then turn to what historians and philosophers of science have contributed to the debate, claiming that the contribution from these science studies scholars has been simply to repeat the arguments made either by Lewontin and Layzer *for* the importance of G×E or by Jensen *against* the importance of G×E. In contrast to these arguments from authority, I attempt to explicate the debate by noting Lewontin and Layzer's employment of $G \times E_D$ and Jensen's employment of $G \times E_B$.

1.4. Difference Mechanisms and the Interdependence of Difference-Makers in Development that Take Different Values in the Natural World

Chapters 2 and 3 set the historical stage, situating the series of debates over G×E in the broader nature-nurture debate between scientists in the biometric research tradition and scientists in the developmental research tradition. The purpose of those chapters is to disentangle the components of the separate traditions as well as the separate concepts of G×E in those traditions. Chapters 4 and 5 re-integrate the components and the concepts, drawing on the philosophical literature on causal-mechanical explanation. In Chapter 4 I take a step back from the particular debates over G×E in order to consider a more general issue in the philosophy of science: causal-mechanical explanation. Philosophers of science in recent years have found a renewed interest in mechanisms. In Chapter 4, however, I argue that these philosophers have been developing a model of causal-mechanical explanation which focuses solely on capturing *regularity*, to the neglect of capturing *variation*. That is, they have been focused on providing an account of causal-mechanical explanation which captures the elucidation of the *regular causal mechanisms responsible for individual development*. But they have failed to consider what role is played by

the *causes of variation responsible for individual* differences in this process of formulating causal-mechanical explanations. This divide obviously lends itself to considering the nature-nurture debate, since it is precisely this divide that has separated the biometric and the developmental research traditions. In Chapter 4 I bridge the divide between the regular causal mechanisms responsible for individual development and the causes of variation responsible for individual differences with the concept of *difference mechanisms*. In short, the thesis is the following: Individual differences are the effect of *difference-makers in development that take different values in the natural world*, and the difference-making variables in the regular causal mechanisms responsible for individual development simultaneously are the causes of variation when the difference-making variables take different values in the natural world. Ultimately, then, the product is a modified account of causal-mechanical explanation that captures both regularity and variation, and which may be utilized to integrate the biometric and the developmental research traditions.

In Chapter 5 I take this general framework between the regular causal mechanisms responsible for individual development and the causes of variation responsible for individual differences and apply it to the conceptual divide between $G\times E_B$ and $G\times E_D$. G×E, I argue, results from the interdependence of difference-makers in development that take different values in the natural world. I utilize this interdependent-difference-makers concept of G×E to integrate $G\times E_B$ and $G\times E_D$. The thesis of Chapter 5 is the following: G×E results from differences in unique, developmental combinations of genotype and environment when both variables are difference-makers in development that naturally take different values and the difference that each variable makes is itself dependent upon the difference made by the other variable; a breakdown in additivity between main effects is a measure of this interdependence of difference-makers. Or, in

other words, the interdependent-difference-makers concept of G×E is just a more general, causal-mechanical reinterpretation of $G \times E_D$, of which $G \times E_B$ is a statistical measure.

2. R. A. FISHER, LANCELOT HOGBEN, AND THE ORIGIN*(S)* OF GENOTYPE-ENVIRONMENT INTERACTION

Abstract. This chapter examines the origin*(s)* of genotype-environment interaction, or G×E. "Origin*(s)*" and not "the origin" because it will be argued that there were actually two distinct concepts of G×E at this beginning: a *biometric* concept, or $G \times E_B$, and a *developmental* concept, or $G \times E_D$. R. A. Fisher, one of the founders of population genetics and the creator of the statistical analysis of variance, introduced the biometric concept as he attempted to resolve one of the main problems in the biometric tradition of biology—partitioning the relative contributions of nature and nurture responsible for individuals differences in a population. Lancelot Hogben, an experimental embryologist and also a statistician, introduced the developmental concept as he attempted to resolve one of the main problems in the developmental tradition of biology—determining the role that developmental relationships between genotype and environment played in the generation of variation. Fisher and Hogben's separate routes to their respective concepts of G×E are outlined, and then these separate interpretations of G×E are drawn on to explicate a debate between Fisher and Hogben over the importance of G×E, the first installment of a debate that still unfolds today. Finally, Fisher's $G \times E_B$ and Hogben's $G \times E_D$ are traced beyond their own work into mid-20th C. population and developmental genetics.

2.1. Introduction

This chapter examines the origin*(s)* of the concept of genotype-environment interaction, or G×E. "Origin*(s)*," and not "the origin," because it will be argued that British biologists and statisticians R. A. Fisher and Lancelot Hogben actually came to consider the concept by quite distinct routes. Fisher, working in the biometric tradition of biology, began by searching for accurate ways to assess the relative importance of nature and nurture; in developing methodologies for the task, he recognized that genotype-environment interactions (or, as Fisher called them, "non-linear interactions") created a potential complication for such assessments. Hogben, working in the developmental tradition of biology, began by evaluating different sources of variability in a population; while he recognized the widely emphasized genetic and environmental sources of variability, he also drew attention to a *third class of variability*: that which arises from the combination of a particular genetic constitution with a particular kind of environment. For

Hogben, this third class of variability was inherently developmental in nature. These distinct routes in these distinct research traditions ultimately led Fisher and Hogben to distinct concepts of genotype-environment interaction. Fisher introduced what will be called the *biometric* concept of G×E, or $G\times E_B$, while Hogben introduced what will be called the *developmental* concept of G×E, or $G\times E_D$. Finally, these distinct concepts led Fisher and Hogben to distinct conclusions when considering the consequences of genotype-environment interactions for assessments of variation in populations. Fisher took the non-linear interactions to be of potential, but unproved, importance; Hogben claimed that they were standard and fundamentally important for understanding variability.

In section 2.2, Fisher's route to $G\times E_B$ within the biometric tradition is traced. It will be seen that his consideration of genotype-environment interaction was a by-product of his developing appreciation for the potential importance of environmental sources of variation along with his development of biometric techniques for assessing such variation. Hogben's route to $G\times E_D$ within the developmental tradition is then taken up in section 2.3. After a brief biographical introduction, Hogben's consideration of genotype-environment interaction is examined, where it will be seen that his interest in the concept emerged out of an earlier appreciation for experimental embryology. In section 2.4, Fisher and Hogben's opposing positions on the importance of genotype-environment interaction are compared. Here the focus will be on revealing how their distinct routes to G×E and the resulting distinct concepts of G×E contributed to their distinct positions when it came to the question of importance. Finally, in section 2.5, the legacies of Fisher's $G\times E_B$ and Hogben's $G\times E_D$ will be traced beyond their own work, acting as a transition into the next chapter.

2.2. R. A. Fisher and the "Non-linear Interaction of Heredity and Environment"

Ronald Aylmer Fisher (1890-1962) looms in the history of 20th C. biology and statistics (Figure 2). His contributions to population genetics, experimental design, significance tests, and general epidemiological methodologies combined with his ardent and infamous endorsement of eugenics, to create a scientist who both revolutionized the biological and statistical sciences, and also vigorously pursued the social and political implications of that revolution.[8] Because Fisher's biography and his contributions to biology and statistics have already been closely examined by historians, philosophers, and sociologists of science, the goal of this section will not be to rewrite this history. Rather, the focus here will be on tracing Fisher's path to genotype-environment interaction, a previously unexamined story. The aforementioned histories, however, will be drawn on quite heavily to reveal how Fisher's attention to genotype-environment interaction was situated within his larger biometric and eugenic research, since the concept was related to each of these domains.

[8] The most complete biography of Fisher comes from his daughter, Box (1978); however, shorter treatments can be found in Mahalanobis (1964) and Yates and Mather (1963). Fisher's contribution to population genetics can be found in Mather (1964), Plutynski (2006), Provine (2001), Skipper (2002), and Thompson (1990). His work on the design of experiments is discussed in Preece (1990) and Yates (1964), and his "logic" of significance tests is examined in Johnstone (1987). Fisher's development of now-classical statistical methodologies is discussed in Anderson (1996), Bennett (1990), Cochran (1980), Finney (1964), MacKenzie (1981), and Rao (1964). The relationship between these biological/statistical contributions and Fisher's interest in eugenics is examined in, for example, Bennett (1983), Kevles (1995), Ludmerer (1972), Mackenzie (1981), Mazumdar (1992), and Soloway (1990). Much of Fisher's scientific correspondence along with Bennett's (1983, 1990) volumes are now available online at the University of Adelaide Library's website as a part of the R. A. Fisher Digital Archive: http://www.library.adelaide.edu.au/digitised/fisher/.

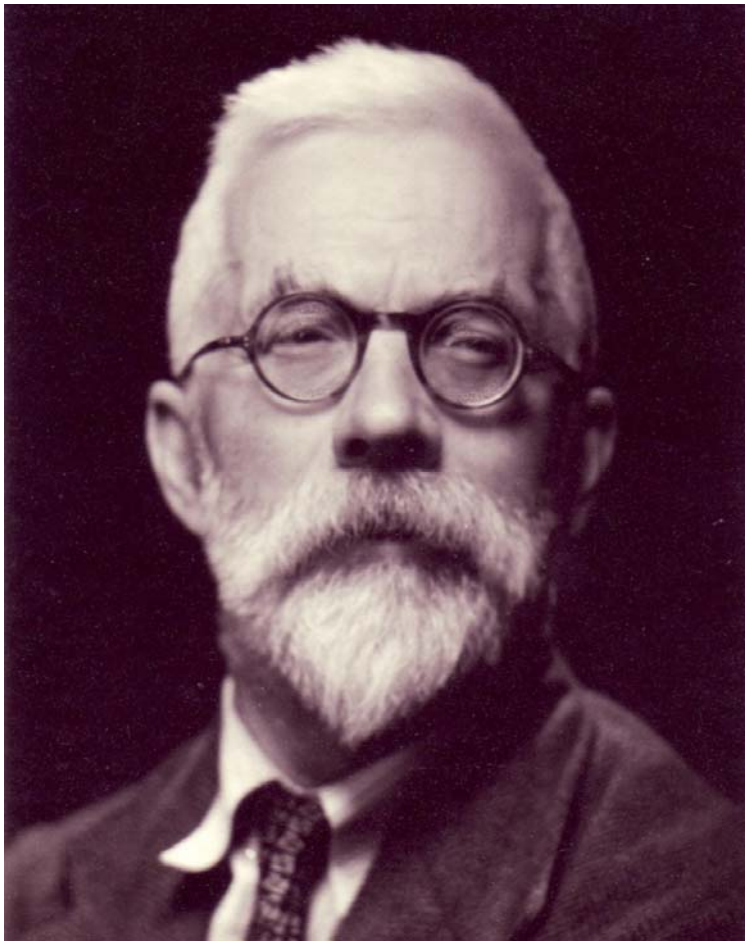

Figure 2. R. A. Fisher. Fisher Papers, Barr Smith Library, University of Adelaide Library, MSS 0013/Series 25. Reprinted with the permission of the University of Adelaide Library.

2.2.1. The Environment Expunged

In October 1918, at only 28 years of age, Fisher published "The Correlation between Relatives on the Supposition of Mendelian Inheritance."[9] Fisher's project was the resolution of the supposed incompatibility between the biometrical theory of continuous variation and the Mendelian theory of discontinuous variation (Norton 1978; Provine 1971).[10] Biometrician

[9] For a commentary on Fisher (1918), see Moran and Smith (1966).
[10] Fisher had actually begun considering the relationship between biometry and Mendelism as early as 1911, when he presented a paper before his Cambridge University Eugenics Society entitled, "Heredity, Comparing the Methods

George Udny Yule, 16 years earlier, had considered the same problem and argued that the Mendelian principles of inheritance could be seen as a special case of the biometric law of ancestral heredity (Yule 1902); Fisher, in contrast to Yule, took the reductive relationship between the Mendelian principles and the biometric law of ancestral heredity in the opposite direction (Tabery 2004). Fisher instead concluded that he came upon "the Law of Ancestral Heredity as a necessary consequence of the factorial mode of inheritance." (Fisher 1918, 421; Sarkar 1998, 106)

But assessing the relationship between biometry and Mendelism was not the only feat accomplished in Fisher's 1918. In the process of deriving the mathematical relationship between the Mendelian principles and the law of ancestral heredity, Fisher also introduced a new statistical concept—*variance* (Box 1978, 53). Fisher was interested in accounting for the sources of variation in a population. Traditionally, populations were statistically evaluated solely with an eye towards averages, but averages shed no light on variation. Fisher noted, though, that if a trait under investigation, such as stature in humans, manifested itself in a population with a normal distribution, then the mean could be calculated along with the standard deviation. Fisher's novel contribution to the statistical analysis of variation in a population was to go beyond the standard deviation and analyze the *square* of the standard deviation:

> When there are two independent causes of variability capable of producing in an otherwise uniform population distributions with standard deviations σ_1 and σ_2, it is found that the distribution, when both causes act together, has a standard deviation $\sqrt{(\sigma_1^2 + \sigma_2^2)}$. It is therefore desirable in analyzing the causes of variability to deal with the square of the standard deviation as the measure of variability. We shall term this quantity the Variance of the normal population to which it refers, and we may now ascribe to the

---

of Biometry and Mendelism" (Fisher 1911). Fisher's 1911 is reproduced along with a discussion of the refereeing of his 1918 in Norton and Pearson (1976).

> constituent causes fractions or percentages of the total variance which they together produce (Fisher 1918, 399).

Six years later, Fisher delivered two lectures to the London School of Economics, in which he explicated the implications of his 1918 essay with data derived by biometricians Karl Pearson and Alice Lee earlier that century (Fisher 1924; Pearson and Lee 1903). Fisher, adding to the discussion of variance, explained, "The amount of variation may be measured either by the Standard Deviation, or by its square, the Variance. When we come to consider the causes of variation, the latter provides the more useful measure. For this reason, that when two independent causes are at work causing variation, the total variance produced is simply the sum of variances produced by the two causes acting separately." (Fisher 1924, 192) Displaying the concept of variance in action, Fisher continued,

> For example, one of the causes of differences in stature is difference of ancestry, the remainder of the causes of variation in stature are those causes which produce variation in stature among girls with the same ancestry, in fact which cause differences in stature between sisters. From measurements of pairs of sisters it is possible to divide up the total variance into two parts. One part representing the differences due to ancestry, the other part representing the other group of causes. If, then, we use the variance as the measure of variability, we can use it to analyse out the fractions of the variability due to different causes; whereas using the standard deviation no such analysis is possible (ibid).

The earlier generation of biometricians, such as Pearson and Yule, had already introduced the concept of the *correlation coefficient* as a numerical measure of association (Mackenzie 1981; Norton 1975). In fact, Fisher worked with Pearson several years before the publication of "The Correlations between Relatives…" solving the problem of deriving the exact distribution of correlation coefficient values in samples from an indefinitely large population (Fisher 1915; Provine 1971).[11] Thus, correlation tables were, by 1918, common; and parental correlations (the

[11] Correspondence between Fisher and Pearson from this period can be found in Pearson (1968).

correlation coefficient comparing a parent with an offspring) along with fraternal correlations (the correlation coefficient comparing two siblings) were frequently calculated from these correlation tables by the biometricians. Fisher employed this correlation technique for partitioning sources of variance in 1918 as a means towards assessing the relative importance of heritable and non-heritable sources of variation, explaining,

> In a similar way each of the ancestors makes an independent contribution, but the total amount of variance to be ascribed to the measurements of ancestors, including parents, cannot greatly exceed one half of the total. We may know this by considering the difference between brothers of the same fraternity; of these the whole ancestry is identical, so that we may expect them to resemble one another rather more than persons whose ancestry, identical in respect of height, consists of different persons. For stature the coefficient of correlation between brothers is about .54, which we may interpret by saying that 54 per cent. of their variance is accounted for by ancestry alone, and that 46 per cent. must have some other explanation (Fisher 1918, 400).

To what cause should this remainder of the total variance be attributed? Perhaps environmental variation? No! Fisher, in 1918, was quick to eliminate this possibility from the minds of his readers: "It is not sufficient to ascribe this last residue to the effects of environment. Numerous investigations by GALTON and PEARSON have shown that all measurable environment has much less effect on such measurements as stature." (ibid) So with environmental variation expunged from the list of possible causes of variation, Fisher had to find another explanation for the 46 percent of the total variance left unaccounted for by ancestry. Fisher responded, "The simplest hypothesis, and the one which we shall examine, is that such features as stature are determined by a large number of Mendelian factors, and that the large variance among children of the same parents is due to the segregation of those factors in respect to which the parents are heterozygous." (Fisher 1918, 400) Fisher continued to draw on the data collected by Pearson and Lee, utilizing their parental and fraternal correlations for stature, span (distance between

fingertips of outstretched arms), and cubit (length of the forearm). He then calculated the variance between siblings attributable to Mendelian segregation and the effects of dominance. With variances due to ancestry, segregation (½ τ^2), and dominance (¾ ε^2) all accounted for, Fisher could finally sum up the sources of the total variance:

Ancestry . . . . 54 per cent.
Variance of sibship:
½ τ^2 . . . . . 31 per cent.
¾ ε^2 . . . . . 15 "
Other causes . . . ---
_____ 46 "
_____
100 per cent. (Fisher 1918, 424)

Fisher famously concluded, "it is very unlikely that so much as 5 per cent. of the total variance is due to causes not heritable, especially as every irregularity of inheritance would, in the above analysis, appear as such a cause." (ibid)

2.2.2. Rothamsted and the Environment Reconsidered—the Origin of $G \times E_B$

Ending an assessment of Fisher's evaluation of the relationship between heredity and environment in the causes of variation, though, would be incomplete if it terminated with his conclusions made in 1918. Historians of genetics and eugenics have often characterized Fisher as a "reformed" or a "new" eugenicist, emphasizing his ultimate recognition of the potential importance of environmental causes of variation (Allen 1986; Kevles 1995; Mazumdar 1992; Soloway 1990).[12] Pauline Mazumdar, in particular, detailed the evolution in Fisher's understanding of the environment's role in variation in her history of the British Eugenics Society (Mazumdar 1992). According to Mazumdar, Fisher's 1918 was, from the very

[12] Barkan (1991) provides a useful comparison of the relationship between "reformed" and "new" eugenics.

beginning, designed to accommodate the ideals of the Eugenics Society: (a) the compatibility of biometry and Mendelism, and (b) the negligible importance of environmental causes of variation (ibid, 110). But in 1919, Fisher left Cambridge and the "loving pressure of the eugenists" to join the Rothamsted Agricultural Research Station as a statistician employed to investigate the effects of environmental variables on crop yield (ibid, 124).[13] At Rothamsted, Fisher was forced to examine environmental variation rather than assume it to be a randomly distributed variable, as he had in his 1918 (ibid, 121).

In 1918, Fisher explained that sources of variation could be summed as long as the causes of variability were *independent*: "…when two independent causes are at work causing variation, the total variance produced is simply the sum of variances produced by the two causes acting separately." (Fisher 1924, 192) Prior to undertaking the work at Rothamsted, the environment could be treated as independent for the simple reason that Fisher took it to be negligible. In making *no* contribution to variability, there was no need for Fisher to concern himself with how environmental variation might be causally related to the other sources of variation. But the research at Rothamsted forced Fisher to reconsider the environment as a possible source of variation. With the environment now on the list of possible sources of variation, Fisher had to also consider the relationship between environmental variation and heritable variation. He judged this possible complication in the second installment of his "Studies in Crop Variation" series, published with W. A. Mackenzie in 1923. He began by warning, "…if important differences exist in the manurial response of varieties a great complication is introduced into both variety and manurial tests; and the practical application of the results of past tests becomes attended with considerable hazard." (Fisher and Mackenzie 1923, 311) The possible difference in manurial response was the possible presence of genotype-environment interaction. "Only if such

[13] Fisher's time at Rothamsted is also discussed in Box (1978, chapter 4) and Mackenzie (1981, chapter 8).

differences are non-existent, or quite unimportant," Fisher continued, "can variety tests conducted with a single manurial treatment give conclusive evidence as to the relative value of different varieties, or manurial tests conducted with a single variety give conclusive evidence as to the relative value of different manures." (ibid) Fisher, here, was making explicit the implications that genotype-environment interaction had on the evaluation of group differences: if genotype-environment interaction existed for a trait under investigation, then examining several varieties' values in just one environment ("a single manurial treatment") would not give conclusive evidence for the relative values of those different varieties in untested environments.

To test for this interaction, Fisher examined the manurial responses of twelve different potato varieties. A relatively small field (0.162 acres) was first divided into two equal parts, one part receiving a farmyard manurial treatment while the other receiving no treatment (see Figure 3). Each half was then itself divided into 36 plots, and each of the twelve potato varieties was planted in triplicate in a chessboard arrangement within each field. Finally, each individual plot was divided again, so that three rows of seven plants were set in each plot; one row received only the basal manuring (B) of the series to which it belonged, while the other two rows received in addition either a dressing of sulphate of potash (S) or a dressing of muriate (chloride) of potash (C).

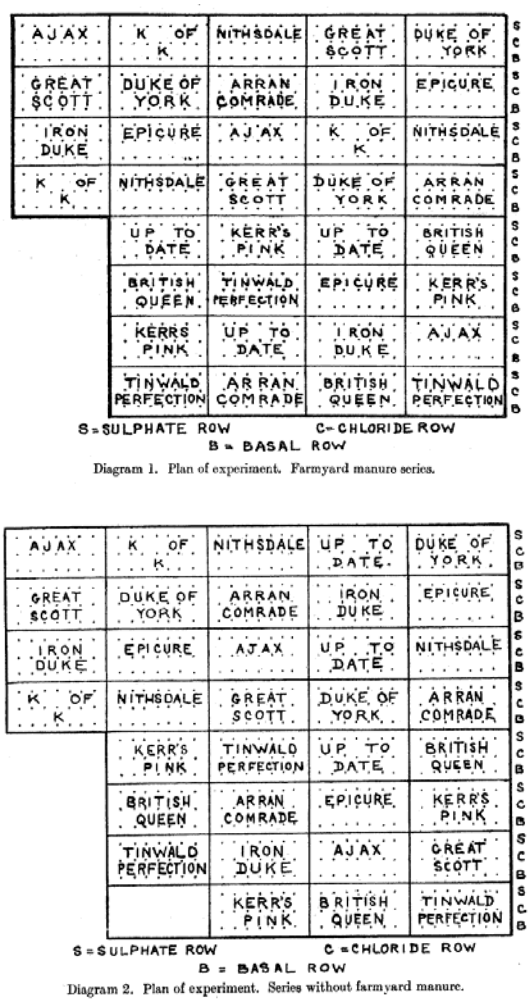

Figure 3. Fisher's experimental design for differential response of potato varieties to manurial treatments. From Fisher and Mackenzie (1923, Diagrams 1 and 2). Reprinted with the permission of Cambridge University Press.

With this experimental design, Fisher was able to use his data to undertake one of his very first applications of an analysis of variance, now a standard resource in any statistician's toolbox. He measured the weight of produce lifted from each of the rows, determining both the mean yield of each of the twelve varieties irrespective of the manuring applied, and the mean yield of each of the manurial treatments irrespective of the variety grown. And what followed

was, as Box (1978, 109-112) has pointed out, the first presentation of the familiar analysis of variance table:

Variation due to	Degrees of freedom	Sum of squares	Mean square	Standard deviation
Manuring	5	6,158	1231·6	35·09
Variety	11	2,843	258·5	16·07
Deviations from summation formula	55	981	17·84	4·22
Variation between parallel plots ...	141	1,758	12·47	3·53
Total	212	11,740	—	—

Figure 4. Fisher's analysis of variation due to manuring, variety, deviations from summation formula, and variation between parallel plots. From Fisher and Mackenzie (1923, Table III). Reprinted with the permission of Cambridge University Press.

The "Deviations from summation formula" category was the measure of the differences between the potato varieties in their manurial response—the measure of interaction. In yet another innovative leap in this same article, Fisher noted that the deviations from the summation formula were not significantly greater than would occur by chance, leading him to conclude, "In the present material evidently the varieties show no difference in their reaction to different manurial conditions." (ibid, 317) This comparison of a source of variation against chance was an early statistical test of significance, or what is now called an "*F*-test" in honor of Fisher's development of the method. Fisher evidently took the results of "Studies in Crop Variation, II" to be quite conclusive. Two years later, in his extremely influential *Statistical Methods for Research Workers* (1925), Fisher again warned of the "interaction of causes" when he introduced the analysis of variance. However, he again used the potato variety-manurial response results from "Studies in Crop Variation, II" to introduce the possible complication and again came to the same conclusion: "There is no sign of differential response among the varieties…" (ibid, 209)

We are now in a position to reflect on Fisher's route to genotype-environment interaction. For Fisher, the concept of interaction was situated in his larger biometric program devoted to measuring the relative contributions of nature and nurture to individual differences in populations, a program initiated by Fisher's mentor and eventual rival, Pearson, the founder of biometry (Porter 2004). Fisher, operating in this biometric tradition, was focused on the causes of variation responsible for individual differences in populations. He asked questions about how much of the variation in populations could be attributed to differences in heredity or differences in environment, and he then developed many of the now-standard statistical methodologies designed to answer these questions, such as the analysis of variance and the statistical test of significance. Non-additive interactions potentially posed a complication for Fisher's summing of variances. But this interaction was understood to be (and, in fact, was defined as) simply a statistical measure—a deviation from the summation formula—which would be detected by Fisher's methodologies if it existed. This notion will be called the *biometric* concept of genotype-environment interaction, or $G \times E_B$. It may be defined as *a statistical measure of the breakdown in additivity between genotypic and environmental sources of variation, which is generated by a statistical methodology such as the analysis of variance* (Tabery Forthcoming).

We are also now in a position to take stock of the various components of the biometric tradition, in which Fisher was operating. Thomas Kuhn's (1962) *The Structure of Scientific Revolutions* forced a generation of historians and philosophers of science to wrestle with how scientific disciplines changed over time, especially in cases of apparent revolution. But to determine how a scientific discipline changed over time, these historians and philosophers of science were required to first determine what exactly a scientific discipline *was*. The result was a bounty of such characterizations: Kuhn's paradigms, Imre Lakatos' research programmes (1970),

Dudley Shapere's domains (1977), Stephen Toulmin's intellectual ecologies (1972), Lindley Darden and Nancy Maull's fields (1977), and Larry Laudan's research traditions (1977).

For my part, I am not so much interested in how Fisher's biometric tradition and Hogben's developmental tradition changed over time as I am interested in the *differences between* these traditions and how the different concepts of genotype-environment interaction developed and persisted in those separate traditions. As a result, I seek only a framework that allows for the articulation of the various components in each tradition so as to compare and contrast them. Probably any of the above frameworks would suffice, although each would, by design, highlight a different facet of the traditions. Lakatos' research programme, for example, would stress the hard core of each tradition along with the protective belt of auxiliary hypotheses, while Toulmin's intellectual ecology would stress the selective forces that drive change for a tradition. However, as my choice of terminology so far will have revealed, I will utilize Laudan's notion of a research tradition. Laudan's framework is suitable for several reasons. He begins by situating a particular research tradition around a particular *problem*. Science, Laudan pointed out, was fundamentally a problem-solving activity, and so understanding a particular research tradition began with identifying the problem on which that tradition was focused. This problem then specified the metaphysical and methodological commitments of members of the research tradition; it established the appropriate entities and processes to investigate, the appropriate questions to ask about those entities and processes, and the appropriate methodologies to employ in seeking to answer those questions. A research tradition, Laudan explained, "is a set of general assumptions about the entities and processes in a domain of study, and about the appropriate methods to be used for investigating the problems and constructing the theories in that domain." (Laudan 1977, 81)

Laudan's framework may be applied to the case of the biometric tradition. The main problem on which Fisher was focused was the partitioning of the *relative contributions of nature and nurture responsible for individual differences* in populations. His approach to causation involved an investigation into the *causes of variation* responsible for these individual differences. He asked, *how much* of the variation in a particular population was due to individual differences in heredity or environment? And he sought to answer those questions with his population-level, *statistical methodologies*. Fisher's route to genotype-environment interaction was in this biometric tradition, and his biometric concept of interaction—$G \times E_B$—bore the marks of that history. The various components of the biometric tradition, now teased apart, are organized in Table 2.

Components	**Biometric Research Tradition**
Problem	Individual Differences
Approach to Causation	Causes of Variation
Causal Question	How Much?
Methodology	Statistical
Concept of Interaction	Biometric—$G \times E_B$

Table 2. The components of the biometric research tradition.

2.3. Lancelot Hogben and the "Interdependence of Nature and Nurture"

In contrast to Fisher, whose name is known to any historian, philosopher, or sociologist of biology, Hogben has received much less attention from those in science studies (Figure 5). As a result, it will be useful to pause before examining Hogben's discussion of genotype-environment interaction and examine the man himself. His sarcasm, quick temper, and tendency to enter public disputes all combined to generate a scientist whose personality, like Fisher's, was just as

large as his scientific pursuits. Moreover, those scientific pursuits were considered throughout much of the early 20th C. to be on par with the contributions of contemporary biologists who are now considered more notable, such as Fisher, J. B. S. Haldane, and Julian Huxley. Influential geneticist C. D. Darlington, as just one instance, wrote of Hogben after his death, "When I was very young, Galdane, Guxley, and Gogben (as the Russians called them), seemed to be the three Magi."[14] (Tabery 2006)

Figure 5. Lancelot Thomas Hogben, Hogben Papers, Special Collections, University of Birmingham Library. Reprinted with the Permission of the University of Birmingham Library.

[14] Darlington to Wells, 6 June 1976, Lancelot Hogben Papers (A.44), University of Birmingham Library. Quoted with the permission of the University of Birmingham Library and P.D.A. Harvey.

Lancelot Thomas Hogben (1895-1975) described his "larval existence" like that of many prominent biologists: obsessively collecting newts, beetles, and butterflies; identifying birds and recognizing them by their eggs; and exploring local geography. "I wanted to be a biologist long before I was twelve," he recalled 60 years later (Hogben 1998[15], 2). Biology, however, was not what God had intended for Hogben…at least that was how his mother saw it. He was born two months prematurely, and to ensure that he would survive the ordeal, his mother dedicated him from birth to the mission field (ibid, 1). This religious devotion was no less powerful on the paternal side of his parenting. Thomas Hogben[16], a self-employed Methodist preacher, spent his days ministering to seamen at the local port under a banner extolling the benefits and burdens of the Christian God: "In the foreground was the lake of brimstone and fire. Across the middle was the edge of a cliff where stood the theatre, the brothel, the casino, the racecourse, the tavern, the *Palais de Danse* and other haunts of Satan. From the edge of the cliff the lost departed were falling in different stages of incandescence. Above the cliff was a solitary pilgrim pursuing a winding road to the rising sun; and, ironically, below it across the flames the legend: *God is Love*." (ibid, 4) Fortunately, the young Hogben and his parents were able to reach a compromise during these formative years; the field of medicine allotted the boy the time to study biology while also preparing himself for service as a medical missionary (ibid, 13).

[15] Hogben wrote his autobiography, *Look Back with Laughter* (*LBL*), in the early 1970's. G. P. Wells (H. G. Wells' son) drew on much of this to write his biographical memoir of Hogben, as a Fellow of the Royal Society (Wells 1978). Wells also edited Hogben's *LBL* with an eye towards publishing it in the late 1970's but could not succeed in the endeavor (Tabery 2006). More recently, Hogben's son and his daughter-in-law have published a *heavily* edited version of *LBL*, under the title, *Lancelot Hogben, Scientific Humanist: An Unauthorized Autobiography* (Hogben 1998). More limited biographies of Hogben can be found in Gurdon and Hopwood (2000), Kevles (1995), Mazumdar (1992), Sarkar (1996), Tabery (2006), and Werskey (1978). For the purposes of this essay, biographical references will be made to Hogben (1998); when material is to be cited that was edited out of Hogben (1998), references will be to Wells' edited version of *LBL* held at the University of Birmingham Library (listed as A.9 and A.10 of the Hogben Papers).

[16] Wells wrote a follow-up essay to his biographical memoir entitled, "Father and Son" (A.38), which detailed the Hogben family along with Thomas Hogben's influences on his son; however, Wells could not convince the Royal Society to publish the sequel.

Largely self-educated at the Stoke Newington Public Library, Hogben excelled academically and won a Major Entrance Scholarship to attend Trinity College Cambridge in 1913 (ibid, 24-25). At Cambridge, Hogben cultivated his biological interests and replaced his parents' religious teaching with a devotion to socialism. He studied botany, physiology, and zoology (winning the Frank Smart Prize for the last in 1915), and also embryonic development at the Marine Biological Laboratory in Plymouth (ibid, 40-41). Hogben entered social life with an equal vigor. Assessing the social societies available to him at the time, Hogben recalled, "I still regard the Union Debating Society of Cambridge (even more that of Oxford) as a potting shed for the cultivation of mentally retarded politicians. The most lively discussions at an intellectually high level were those which took place at the *Moral Sciences*, colloquially *Moral Stinks, Club*, where Bertrand Russell and [G. E.] Moore minced words with their philosophical competitors, in the Fabian Society and its study circles, and in the *Heretics* founded by C. K. Ogden of Basic English fame." (ibid, 33) The Fabian Society at Cambridge was a particularly accommodating match for Hogben; he met his first wife, Enid Charles, there and eventually became its secretary, changing the Society's name to the University Socialist Society (ibid, 51).

At the outset of World War I, Hogben joined noncombatant Quaker relief organizations—first the War Victims contingent, which helped house French civilians rendered homeless by the combat, and then the Friends' Ambulance Unit (ibid, 48-49). When the British government introduced compulsory military service, though, in 1916, Hogben protested this action as a conscientious objector and spent several months imprisoned in Wormwood Scrubs for the decision (ibid, chapter 7). After the War, Hogben entered the academic life, teaching and leading research in London at Birkbeck and the Royal College of Science (1917-1922), in Edinburgh at the Animal Breeding Research Laboratory (1922-1925), in Montreal at McGill

University (1925-1927), at the University of Cape Town (1927-1930), at the London School of Economics (1930-1937), at the University of Aberdeen (1937-1941), at the University of Birmingham (1942-1943), in London at the War Office (1944-1946), and finally at the University of Birmingham again where he retired (1947-1961).

In his early career at Birkbeck, Edinburgh, McGill, and Cape Town, Hogben was primarily devoted to experimental embryology and physiology. He worked on the mechanisms of amphibian metamorphosis with Julian Huxley and on the amphibian pigmentory effector system with Frank Winton (Hogben and Winton 1922a, 1922b, 1923; Huxley and Hogben 1922). The investigations were largely interventionist by nature; for example, he isolated the role of the pituitary in the pigmentory effector system by surgically going through the roofs of frogs' mouths and removing various portions of the gland, then noting the subsequent lack of pigmentation (see Figure 6). Hogben, while at Edinburgh and with the help of Huxley and Frank A. E. Crews, also founded the Society for Experimental Biology and its accompanying *British Journal of Experimental Biology* (Crews et al. 1923; Erlingsson 2005; Hogben 1998, 79).

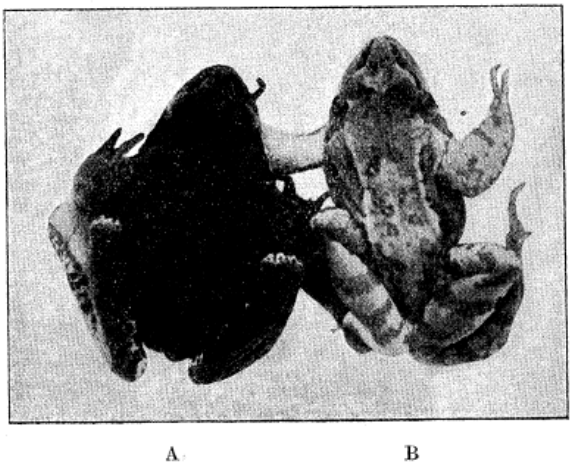

Figure 6. Two frogs, 19 days after pituitary operation by Hogben. (A) partial removal of only anterior lobe. (B) complete removal. From Hogben and Winton (1923, Figure 2). Reprinted with the permission of The Royal Society.

It was Hogben's 7 years at the LSE, however, that produced his most lasting contributions to science and society. During these years he wrote his first two, hugely successful, Primers for the Age of Plenty: *Mathematics for the Million* (1937) and *Science for the Citizen* (1938). Though Hogben clearly relished these exercises in popular science, he was also wary of their impact on his reputation as a respectable scientist. Hogben wrote *Mathematics for the Million* in 1936; however, he held off on publishing it with his name as the author because "At that time, [he] was still a candidate for the Fellowship of the Royal Society, and its hierarchy frowned formidably on what they regarded as scientific popularisation." (Hogben 1998, 138) The manuscript was thus left in a drawer.[17] Only after becoming a Fellow, Hogben took advantage of a fortuitous conversation with the American publisher, Warder Norton:

> To my query about whether he had any special mission this side of the pond, he replied that there would, in his view, be a big market for a book if it could do for mathematics what [H. G.] Wells had done for world history in his *Outline*. When I asked whether he had any prospective author in mind, he mentioned Bertrand Russell as his best bet. I reminded him that Bertie tells us somewhere how, as a boy, he read through all the books of Euclid in one stride and decided that Euclidean geometry was too easy to merit further study. Besides, I added, Bertie had discovered that most children dislike mathematics. Warder Norton's disappointment was so patent that I said something to the effect that I had already written the book he wanted, and that I might be able to find it for him (ibid, 137).

But even after becoming a Fellow, Hogben felt the need to claim that the manuscript wasted none of his professional time. In the preface, he claimed that he "wrote this book in hospital during a long illness for my own fun." (Hogben 1937, xi) His son, however, remembered differently: "Lancelot repeatedly claimed that he wrote *Mathematics for the Million* in six weeks

[17] Hogben, at this point, actually did offer the manuscript to a publisher free of charge under the condition that it be published anonymously, but the publisher had no interest because "mathematics is a worst seller." (ibid, 138)

while hospitalized for radical (barbaric) surgery on his nasal sinuses. Of course, this was implausible. I still remember sitting by Lancelot with a shawl over his head as he inhaled balsam or some such noxious vapour. He was still working out ideas. I served as a sounding board." (Hogben 1998, xiv)

Hogben, during his stay at the LSE, also attacked Britain's eugenics movement with a tenacity unmatched even by the standards of other anti-eugenicists of his day.[18] Sir (later Lord) William Beveridge, then the director of the LSE, sought to bridge the divide between the natural and the social sciences and so announced the search for a Chair of Social Biology in 1929, which would be funded by the Rockefeller Foundation (Kevles 1995, Mazumdar 1992, Werskey 1978). Fisher applied for the position (Box 1978, 202), but it was Hogben who was ultimately invited to take the post. In his autobiography years later, he recalled this vocational victory with glee, noting, "…the brass hats of the Eugenics Society were already congratulating themselves on the prospect of one of their co-religionists getting the job." (Hogben 1998, 121) Hogben, however, only agreed to take the appointment after some reluctance, later explaining,

> At that time human genetics was a morass of surmise and superstition. It had as yet no sufficient theoretical foundation for firm conclusions about the results of matings necessarily beyond the range of experimental control. In short, no advance could materialise without further mathematical exploration of the postulates of experimentally established principles. At first, I was appalled by the prospect of engaging in a task so formidable, and one for which I could not formulate a programme on the spur of the moment. It was, however, a social as well as a scientific challenge. The rationalisation of race prejudice by appeal to biological principles was then plausible only because human genetics was so immature. Should I prosper in

[18] Hogben's role in the anti-eugenics response to the eugenics movement is discussed in Barkan (1991); Blacker (1952); Kevles (1995); Ludmerer (1972); Mazumdar (1992); Paul (1995, 1998); Soloway (1990); and Werskey (1978). Those historians that consider the origins of Hogben's attention to the role of environmental sources of variation (such as Werskey (1978) and Mazumdar (1992)) will be discussed below in section 2.4 when that topic is addressed.

> the Herculean task of cleaning the Augean stables of human heredity, I should be contributing to the overdue disposal of a manure heap of insanitary superstitions (Hogben 1998, 122).

Ultimately, Hogben recalled, it would be one of his fellow-"Magi" that convinced him to take on the responsibility: "Conversation with J. B. S. Haldane jerked me out of indecision concerning my fitness for the task." (ibid) Hogben accepted the position and left Cape Town, joining the LSE in 1930.[19]

2.3.1. Cleaning the Augean Stables

Hogben's first full-fledged assault on the eugenics movement came with the publication of his *Genetic Principles in Medicine and Social Science* (1932a).[20] "This book does not undertake to set down all that is known and has been surmised about human inheritance," Hogben admitted. Instead, it was the first step in his Herculean task: "It is an attempt to separate the wheat from the tares, to indicate where a sound foundation of accredited data is available, to discuss what methods can be applied to the extremely elusive nature of the material with which the human geneticist deals, and to re-examine some of the biological concepts which have invaded other fields of inquiry in the light of modern advances in experimental genetics." (ibid, 9) Hogben reviewed the problem of twin resemblance, research identifying single gene substitutions related to human pathology, the use of serological data for mapping chromosomes, the genetic basis of social behavior, the concept of a race, the nature of genetic selection in a social group, the growth of human populations, and the social applications of genetic principles. The underlying thread that united all of these discussions was his persistent emphasis on the role that the environment played in the development of pathological, behavioral, and social traits. For

[19] A thorough discussion of Hogben's appointment at the LSE can be found in Mazumdar (1992, chapter 4).
[20] "Full-fledged" because Hogben had addressed critically the science of eugenics to a limited extent prior to *Genetic Principles* in earlier book chapters and lectures (see, for example, Hogben 1927, 1931a, 1931b).

instance, with respect to pathological traits, Hogben emphasized the importance of environmental agencies contributing to "deficiency diseases" such as rickets (ibid, 64). And when arrested social behavior in the case of Mongolism was discussed, Hogben drew attention to the effect of birth order and the uterine environment on the incidences of such a trait (ibid, 99-103).

The role of the environment was of such prominence in the pages of *Genetic Principles* because Hogben felt that biologists had generally learned to neglect it in response to theoretical developments of the previous century. More specifically, the death of Lamarckism, the discovery of cellular fertilization, and finally the rise of Weismann's theory of the germ plasm ushered in a generation of biologists with no theoretical interest in the environment (ibid, 39-40). Hogben, however, also noted that this tendency to ignore the environment was gradually eroding in the face of experimental biology, and especially experimental embryology:

> Weismann's teaching had a profound influence on the form which the hypothesis of natural selection assumed in the closing years of the nineteenth century. It has left a profound impress upon biological discussion of social evolution. During the present century the rise of experimental methods in the study of heredity and development has shown the immense importance of environment in determining individual variability among animals and plants. Strictly speaking, it is meaningless to speak of hereditary characters. Characters as such are the end-product of a prolonged and immensely complex series of reactions between the structural materials contributed by the sperm and the egg on the one hand, and all the characteristics of the physical medium which the cells descended from a given fertilised egg develop (ibid, 40).[21]

[21] With regard to selection, Hogben described a similar situation: "The selection doctrine assumed a more rigid form when it was robbed of the Lamarckian assumptions implicit in much of Darwin's earlier writings. In the hands of Weismann and the Biometricians the implications of "blending" inheritance were more explicitly formulated. Heredity and variation were necessarily co-extensive. Environment as an aspect of the problem of development faded out of the picture. For a generation biologists were hypnotised by the discredit into which the Lamarckian teaching had fallen, till the progress of experimental embryology and the new cell anatomy relegated Weismann's hypothesis of germinal selection to the same limbo as the Lamarckian doctrine. While its influence persisted, all differences between parents and offspring were regarded as genetic." (ibid, 146)

In addition to describing the state of early-20th C. biology, this quote also revealed Hogben's prescription for formulating the relationship between heredity and the environment. "Genetical science," Hogben claimed, "has outgrown the false antithesis between heredity and environment productive of so much futile controversy in the past." (ibid, 201) Since every character is the end-product of an immensely complicated series of reactions between external agencies and the hereditary material, "Differences can be described as determined predominantly by hereditary or predominantly by environmental agencies if, and only if, the conditions of development are specified." (ibid, 98) To drive this point home, Hogben pointed out that variation in a population could arise from hereditary variation (emphasized by eugenicists), environmental variation (emphasized by anti-eugenicists), and an often-ignored *third class of variability*: that which "arises from the combination of a particular hereditary constitution with a particular kind of environment." (ibid, 98) It will be especially important to keep this conception of the relationship between heredity, environment, and development in mind when Hogben employed this relationship to criticize Fisher the following year.

But in 1932, Hogben had not yet criticized Fisher, and in his review[22] of *Genetic Principles* Fisher welcomed Hogben's position at the LSE, despite the fact that his application was turned down in favor of Hogben's. Fisher began, "[Hogben's] recent appointment as Professor of Social Biology at the London School of Economics gave the welcome assurance that his keenly analytic brain, and training in a severe experimental discipline, would be put to important service in the study of the biology of man. The rapid appearance of his new book, 'Genetic Principles in Medicine and Social Science,' will therefore be received with more than ordinary interest by all those who recognize the need, in this field, of whole-time workers with an adequate biological training." (Fisher 1932, 147) Fisher was especially impressed by

[22] For other prominent reviews of Hogben (1932a) by his contemporaries, see Huxley (1932) and Haldane (1932).

Hogben's assimilation of the most recent work in the field and in his "strong taste for analytic precision of statement." (ibid) However, Fisher also worried that this attention to "purely academic considerations" led to an exclusion of "aspects of more practical importance." (ibid) "Throughout the book," Fisher complained, "those who consider that the practical importance of the problem renders it urgent, will receive a disturbing impression that they are being asked to wait, in solemn hush, outside the laboratory door, until the Professor sees fit to announce that the ultimate truth has at last been revealed." (ibid, 147-148)

Fisher also took issue with Hogben's discussion of the environment. With regard to Hogben's account of the source of some biologists' neglect for the role of the environment, Fisher called it a "historical misapprehension" to suggest that Galton and Weismann's rejection of the Lamarckian inheritance of acquired characters "led to a neglect of the somatic importance of such modifications." (ibid, 149) In contrast to Hogben's claim that Galton and Weismann's influence led to conceiving of all differences between parents and offspring as genetic, "It would be truer to say that this was so while the influence of the Lamarckian doctrine persisted, for it was the distinctive dogma of this doctrine that such differences, even if environmentally induced, were inherited. Only when Lamarckism had been overthrown could the problem of the relative importance of the congenital and the induced differences be clearly formulated." (ibid) And when Hogben summarized the connection between mental defect, scholastic success, and birth order by writing, "This connection leaves little doubt that environmental influences play a very significant rôle in determining the manifestation of mental defect" (Hogben 1932a, 106), Fisher concluded, "We can only hope that when Professor Hogben has had sufficient leisure to produce the authoritative work, which we may later hope for, that he will think it better to omit, or radically to rewrite, his discussion of this type of material." (ibid, 150)

2.3.2. The William Withering Memorial Lectures

That same year (1932), the Medical Faculty of the University of Birmingham invited Hogben to deliver their William Withering[23] Memorial Lectures, and Hogben chose medical genetics as the theme of his Lectures (Hogben 1998, 123). Hogben, in preparation for the Lectures, contacted Fisher in February of 1933 on a point of clarification:

> Dear Fisher, I am at present engaged in preparing a course of lectures in which I shall be dealing with your own contributions to the genetic theory of correlation. There is one point in your 1918 paper which worries me very much. When you speak of the contribution of hereditable and nonhereditable causes of variance in a population, what exactly do you mean? I often use the same form of words myself and lately I have been searching for a more explicit formulation of the problem. Suppose you say that 90 per cent of the observed variance is due to heredity, do you mean that the variance would only be reduced ten percent, if the environment were uniform? Do you mean that the variance would be reduced by 90 per cent, if all genetic differences were eliminated? Perhaps you will think the question silly; but if you could suggest an alternative form of words, it might help.[24]

Fisher responded the following day.

> Dear Hogben, Your question is a very sound one. The point is this:-If the differential effects of environment and heredity are not correlated, i.e. if each genotype has an equal chance of experiencing with their proper probabilities, each of the available kinds of environment, then the variance is additive, and the statements you have are equivalent. If they are not independent, then the practical choice of a form of statement will depend upon what the correlation is due to.
>
> E.g. if congenitally browner people expose themselves more to the sun than others, then eliminating the congenital difference will carry with it the elimination of some of the environmental difference, which virtually belongs to it. Equally in this case elimination of environmental differences

[23] William Withering (1741-1799) was a botanist and physician (discovering the medical implications of foxglove's active ingredient—digitalis), a member of the Lunar Society, and a chief physician at the Birmingham General Hospital (Aronson 1985).
[24] Hogben to Fisher, 17 February 1933, R. A. Fisher Papers (available on-line at http://www.library.adelaide.edu.au/digitised/fisher/). Quoted with the permission of Leslie Hogben.

should mean the elimination not only of differences within the same or equivalent genotypes. On the other hand if exposure to sun induced germinal brownness, the other of the two partitions mathematically available, will be the practically useful one.[25]

Fisher took Hogben's question to be one concerning genotype-environment correlation, and so answered Hogben's question with a discussion of a genotype's "chance of experiencing" a particular environment. The concept of genotype-environment correlation refers to cases where an individual's genotype correlates with exposure to particular environments. Genotype-environment correlation, however, was not Hogben's target, and he took several days to construct a lengthy rebuttal. "Dear Fisher, I don't think you quite got the difficulty which I am trying to raise. It concerns an inherent relativity in the concepts of nature and nurture, which did not emerge clearly so long as geneticists drew a hard and fast line between metrical and unit characters."[26] Hogben then introduced an example to clarify his concern.

> Let me take an example which is particularly pregnant because the character can be defined either as an all or none reaction or in metrical terms. I refer to the bar eye series in Drosophila. From Krafka's data you will see the following values for facet number are given at 15° and 25° C.

	Low bar	Ultra bar
15° C	189	52
25° C	74	25

> Consider the elementary population with the following structure. The genotypes are Low bar and ultra bar in equal numbers, equally distributed between two environments, namely an incubator at 15° C and one at 25° C. There is zero correlation between the distribution of environmental and genetic variables. Yet I cannot agree that the two statements "y per cent of the variance is due to environment," and "the variance

[25] Fisher to Hogben, 18 February 1933, R. A. Fisher Papers (available on-line at http://www.library.adelaide.edu.au/digitised/fisher/). Quoted with the permission of the University of Adelaide Library.
[26] Hogben to Fisher, 23 February 1933, R. A. Fisher Papers (Series I, Hogben, L.), University of Adelaide. Quoted with the permission of Leslie Hogben.

> would be reduced by y per cent if all differences of environment were eliminated," are equivalent nor that there is equivalence between the two statements "x per cent of the variance is due to heredity" and "the variance would be reduced by x per cent if there were no genetic differences."

Hogben then pointed out that the result was a "lack of singularity."

> The fact is that there is a lack of singularity in the problem when it is reduced to practical form, as can be seen in arithmetical form in this instance. In the population defined the mean is 85 and the variance is 3906 to the nearest integer. Let us abolish all differences of environment. We can do this in an infinite number of ways. One would be to culture all flies at 15° C. Result: mean 120.5 and variance 4692. Another is to culture them all at 25° C. Result mean 49.5 and variance 600. Which of these two variances has priority as an estimate of the "contribution" of environment to the observed variance in the fourfold population? Again we eliminate all genetic differences by killing off all ultra bar flies. Result: mean 131.5 and variance 3306. We could alternatively kill off all low bar flies. Result: mean 38.5 and variance 182. Which of these gives the contribution of heredity to the observed variance? (ibid)

Hogben shoved aside the matter of genotype-environment correlation here: "There is zero correlation between the distribution of environmental and genetic variables." Instead, he focused on the "lack of singularity" which resulted from a variable response of two genotypes to an array of environments. He closed by reemphasizing his lack of interest in genotype-environment correlation and explaining the motivation behind his interest in the "lack of singularity" problem, concluding, "The point I am after is not what assumptions about the distribution of the environment and the distribution of gene differences are made in the mathematical formulation of the problem. Obviously we can make more or less arbitrary assumptions about that. *What I am worried about is a more intimate sense in which differences of genetic constitution are related to the external situation in the process of development.*" (ibid, emphasis added)

Hogben's letter on February 23rd marked the dawn of genotype-environment interactions being utilized as a critical tool to attack the summing of heritable and non-heritable sources of

variance. The bar-eye *Drosophila* example also became the empirical backbone of Hogben's last William Withering Memorial Lecture, entitled "The Interdependence of Nature and Nurture" in the published form of these Lectures (*Nature and Nurture*, 1933a), and entitled "The Limits of Applicability of Correlation Technique in Human Genetics," published in the *Journal of Genetics* that same year (1933b).[27] It was, in short, an all-out attack on Fisher. Fisher's 1918 was noteworthy for human genetics, Hogben claimed, both because of the "thoroughness with which he assailed the mathematical intricacies of a purely genetical theory of correlation" and also because of the "particular conclusions about nature and nurture advanced in his memoir." (1933a, 92) Hogben reiterated Fisher's (1918) objective: to use fraternal correlations to "determine the respective contributions which nature and nurture make to the variability of a normal population, using the mean square deviation as the measure of variability." (ibid) And Hogben identified the passage from Fisher (1918), which particularly concerned him, quoting, "it is possible to calculate the numerical influence not only of dominance but of the total genetic and non-genetic causes of variability. An examination of the best available figures for human measurements shows that there is little or no indication of non-genetic causes. The closest scrutiny is invited on this point not only on account of the practical importance of the predominant influence of natural inheritance, but because the significance of the fraternal coefficient in this connection has not previously been realised." (ibid; from Fisher (1918, 433))

Because of the centrality that correlation coefficients played in Fisher's 1918, Hogben devoted his essay to critically assessing the correlation technique. Hogben admitted that the technique of correlation could be "used to detect the existence of differences due to environment

[27] Hogben's *Nature and Nurture* (1933a) is much more often cited than his *Journal of Genetics* paper (1933b), so references will be made to his 1933a. However, although the essays are extremely similar, the version that appeared in the *Journal of Genetics* did go into slightly more detail at important points. I will say explicitly when references are being made to those portions uncontained in *Nature and Nurture*.

and differences due to heredity." (ibid, 93) Based on correlation coefficients from monozygotic and dizygotic twins, Hogben even conceded that "Few biologists would hesitate to draw the conclusion that intellectual differences may arise because of gene differences." (ibid) However, moving beyond the detection of such differences, "The difficulties of interpretation begin when we attempt to clarify what is meant by calculating 'the numerical influence…of the total genetic and non-genetic causes of variability.'" (ibid, 94-95) Hogben drew on his Cambridge, philosophical hero to make this point: "In his illuminating essay on the *Notion of Cause* Bertrand Russell has pointed out that few words are used with greater ambiguity in scientific discussion."[28] (ibid, 95) What Hogben had in mind here was an extension of the critique of Galton and biometry he first made in his *Genetic Principles*. "The biometrical treatment of variability," Hogben argued, "inherited from Galton a tradition of discourse in which the ambiguity of the concept of causation completely obscured the basic relativity of nature and nurture. Since then this relativity has become increasingly recognised through experiments involving the use of inbred stocks in physiological laboratories, especially in connexion with experimental work on diet. It is therefore necessary to examine with great care what we mean when we make measurements of a genetic difference and a difference due to environment." (1933a, 95)

To drive home this point, Hogben introduced to his reader the same case he introduced to Fisher in correspondence earlier that year, providing both the data and, this time, a reaction norm graph of the differential responses of low-bar and ultra-bar *Drosophila* strains to a variable environment (see Figure 7).[29] The differences between points A and B ($_{16}\delta_H$)[30], and between

[28] This was not the first time Hogben revealed his philosophical indebtedness to Russell. Hogben's *The Nature of Living Matter* (1931a), a mechanistic critique of vitalism, was dedicated to Russell.
[29] For a history of the reaction norm concept, see Sarkar (1999).

points E and F ($_{25}\delta_H$) corresponded to what Hogben claimed experimental biologists meant by a genetic difference. Meanwhile, the differences between points B and C ($_B\delta_E$), and between points D and E ($_A\delta_E$) corresponded to what Hogben claimed experimental biologists meant by a difference due to environment. Hogben assessed, "Clearly we are on safe ground when we speak of a genetic difference between two groups measured in one and the same environment or in speaking of a difference due to environment when identical stocks are measured under different conditions of development." But then he continued, questioning, "Are we on equally safe ground when we speak of the contribution of heredity and environment to the measurements of genetically different individuals or groups measured in different kinds of environment?" (ibid, 97) Hogben asked his reader to consider a low-bar stock kept at 16° C and an ultra-bar stock kept at 25° C, creating the observed differences AC or DF. "How much of AC or DF is due to heredity and how much to environment? The question is easily seen to be devoid of a definite meaning." (ibid) He then drew for his reader the same conclusion he drew from this data when he wrote to Fisher previously:

> We might be tempted to say that the genetic contribution is the difference which would exist if both stocks had been cultured at the same temperature. This could be done in an infinite number of ways. If they were both cultured at 16° C. heredity would contribute the difference AB. We might be tempted to say that the contribution of environment represents what the difference would be if all the flies belonged to the same stock. Obviously this can be done in at least two ways. Keeping the same difference of environment we might substitute low-bar individuals for the ultra-bar stock at 25° C. The difference between the two sections of the population would then be represented by DE. If we substituted ultra-bar individuals for the low-bar stock at 16° C. the difference would be BC. Either ED or BC is equally entitled to be regarded as the contribution of environment (ibid).

[30] Krafka (1920), as will be discussed below, took measurements at a number of different temperatures including both 15° and 16° C. As a result, it is not necessarily a typographical error that Hogben used 16° C. here but used 15° C. in his correspondence with Fisher.

Hogben importantly emphasized, "The literature of experimental physiology is not wanting in examples of such divergent curves representing the measurement of a character and the strength of the environment." (ibid) He drew on the research of Norman B. Taylor (1931), and Frank R. Winton (1927), his former colleague and co-author from Edinburgh, who respectively examined variation in the sinus beat of *Xenopus* and *Rana* with regards to temperature, and variation in the mortality rate of rats with regards to red squills (Hogben 1933b, 385).[31]

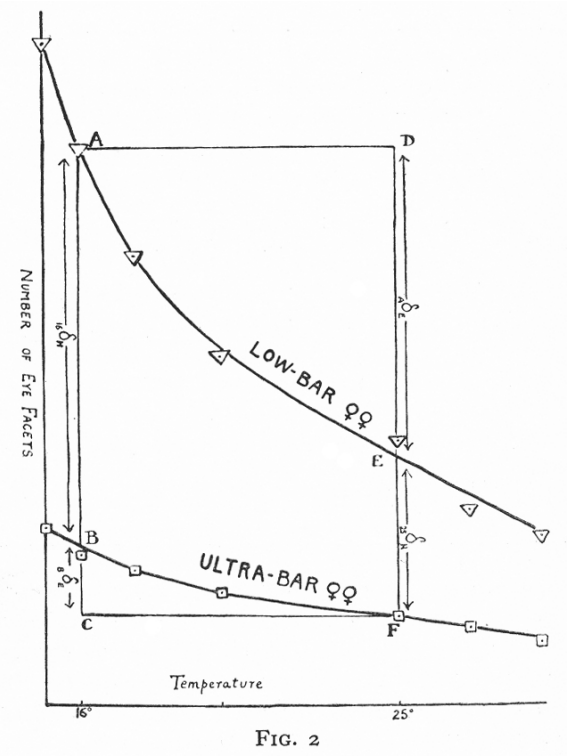

FIG. 2

Figure 7. Hogben's norms of reaction for low-bar and ultra-bar *Drosophila* strains derived from Krafka (1920). Figure from Hogben (1933a, Figure 2). Reprinted with the permission of Macmillan.

[31] These references are not included in the version of this essay which appears in the William Withering Lectures published as *Nature and Nurture* (1933a). There, Hogben only wrote, "There is no reason to multiply instances in order to show the need for extreme care in formulating the problem of nature and nurture in quantitative terms." (Hogben 1933a, 97)

Hogben also explicated the *practical* implications of this variable response by different genotypes to environmental differences. “The only practical significance which Fisher’s analysis of variance seems to admit is that, if it were correct, we could only reduce variance with respect to stature in a human population by 5 per cent. or less if the environment were perfectly uniform.” (Hogben 1933a, 114) As Hogben pointed out above, though, creating such a uniformity can be done in an infinite number of ways, “some tending to bring out genetic differences which were not previously measurable, others tending to obscure genetic differences which were measurable before.” (ibid) Hogben called the calculation devised by Fisher (1918) to add up all the sources of variance a “balance sheet of nature and nurture” (see section 2.1 above). And he asked, “Has a balance sheet of nature and nurture any meaning in this sense, unless we assume that the variance of a population, if affected at all, is necessarily diminished when the environment is made more homogenous?” (ibid) But as he wrote to Fisher before, and as he would repeat in published form, “Such an assumption is certainly false.” (ibid) With regards to Krafka’s two *Drosophila* populations (low-bar and ultra-bar) exposed equally to the two environments (15° and 25° C), the variance was 3906.5. Creating environmental uniformity by confining both stocks solely to 15° C would *increase* the variance to 4692.25. But creating environmental uniformity by confining both stocks solely to 25° C would *decrease* the variance to 600.25. Hogben asked his reader the same question he asked Fisher: “Have either of these estimates any special priority as a measure of the contribution of heredity alone to the observed variance?” (ibid, 116) Likewise, genetic uniformity could be created by substituting low-bar stock for the ultra-bar individuals, generating a variance of 3306.25. Or, genetic uniformity could be created by substituting ultra-bar stock for the low-bar individuals, generating a variance of

182.25. Again, "Which of these two estimates gives the contribution of environment alone?" (ibid) Hogben concluded,

> In whatever sense Fisher himself intended his balance sheet to be interpreted, there is no doubt that many writers on human biology entertain the belief that biometrical estimates of this kind do entitle us to set such limits. On the basis of such statements as the previous quotation about stature, it is often argued that the results of legislation directed to a more equitable distribution of medical care must be small, and that in consequence we must look to selection for any noteworthy improvement in a population. This is rather like saying that the difference between black and white is negligible because an inkpot thrown into a tank of china clay has very little effect on the latter (ibid, 116-117).

We can only assume that Fisher felt little gratitude when Hogben concluded his *Journal of Genetics* essay by writing, "It is a great pleasure to acknowledge the courtesy with which Dr. Fisher has replied to communications in which some of the issues raised in this discussion have been explored." (Hogben 1933b, 405)

2.3.3. From Development to Interaction—the Origin of $G \times E_D$

We have already ascertained how Fisher came to consider the question of genotype-environment interaction (see section 2.2.2 above). In developing the analysis of variance, Fisher recognized quite early on that such non-linear interactions would create complications for assessments of the relative contributions of heritable and non-heritable sources of variation (Fisher and Mackenzie 1923). However, in his own empirical research on potato varieties, Fisher found no such interaction.

But how did Hogben come to consider genotype-environment interactions? It was apparently *not* through a familiarity with Fisher's own research on the topic. Hogben never mentioned Fisher's "Studies in Crop Variation, II" or his discussion of the topic in *Statistical Methods for Research Workers*. More tellingly, Hogben first introduced the problem to Fisher in

correspondence as if it was an issue with which Fisher might have no concern, admitting, "Perhaps you will think the question silly."

Historians who have considered Hogben's criticisms of the eugenicists have tended to explain Hogben's attention to environmental sources of variation by appeal to political motivations. Gary Werskey (1978) and Mazumdar (1992) both pointed to the influence of the Second International Congress on the History of Science, held in 1931 in London, on Hogben and other left-wing British scientists of the day, such as Haldane, Joseph Needham, Hyman Levy, and J. D. Bernal. At this conference, a Soviet delegation led by Nikolai Bukharin introduced Marxism to the British scientific community. Mazumdar explained, "Hogben's thinking on the problems of social biology did not take a completely new direction following his contact with Marxism, but the Marxist analysis both sharpened his perception of the class-bound nature of the eugenic programme, and also provided a theoretical support for his campaign against the over-emphasis of the biological in human society." (ibid, 161) And Werskey wrote, "Rather than completely sacrifice his outside political interests to the demands of scientific life, he consciously brought his politics to bear on the kind of science he did. As a feminist who was also an experimental biologist, Hogben was drawn in the early twenties to the new field of comparative endocrinology, in order to study the hormonal bases of sex differences. As a socialist, he likewise found himself attracted to the social biology of class and racial differences." (ibid, 105)

However, while such political analyses may help to explain Hogben's *motivations* for attacking the eugenics movement, they do little to explain the actual *tools* of the attack itself. This point should not be taken as a criticism of these histories; Mazumdar fully admitted, "Marxism helped Hogben to define his problem, but it did not provide him with the tools with

which to solve it." (ibid) Rather, the point is that a closer analysis of Hogben's actual criticisms requires more than an appeal to his political motivations. For Mazumdar, that closer analysis came from assessing the influence of German mathematical genetics (*Vererbungsmathematik*) on Hogben's subsequent research. A familiarity with the work of Wilhelm Weinberg, Fritz Lenz, and Felix Bernstein, Mazumdar revealed, led Hogben to introduce to the English-speaking world new mathematical techniques for analyzing pedigree data. Hogben, in 1931 and 1932, published in the *Journal of Genetics* a series of papers on "The Genetic Analysis of Family Traits" applying the *Vererbungsmathematik* approach to pedigree analyses of traits caused by single gene substitutions, double gene substitutions, and single recessive genes (Hogben 1931c, 1932b, 1933c; Mazumdar 1992, 162-169).

But Hogben's discussion of genotype-environment interaction was quite distinct from his discussion of pedigree analyses; the latter was a tool used to reform a methodology employed by eugenicists, while the former was a tool used to critically attack eugenic interpretations of variance analyses. As a result, a familiarity with German mathematical genetics will not suffice to explain the origins of Hogben's consideration of genotype-environment interaction. Fortunately, Hogben left a revealing clue for this explanation in the form of his last line to Fisher: "What I am worried about is a more intimate sense in which differences of genetic constitution are related to the external situation in the process of development." Hogben, here, explained quite clearly what motivated his interest in genotype-environment interaction—an appreciation for the developmental relationship between genotype and environment, and the variation that resulted from that relationship.

Considering the developmental relationship between the genotype and the environment was nothing new for Hogben in 1933. We saw above that his earlier *Genetic Principles* was

filled with warnings against only construing phenotypic variation as either a product of genetic differences or of environmental differences. There, Hogben criticized the "false antithesis of heredity and environment" (Hogben 1932a, 201). He admitted that some hereditary variability would exist in almost any environment; and, likewise, some variability would be brought about by the environment acting on the same genetic material. However, Hogben also drew attention to a *third class of variability*, which "arises from the combination of a particular hereditary constitution with a particular kind of environment." (ibid, 98) In 1932, when *Genetic Principles* was published, the only empirical example Hogben gave of this third class of variability came from the "abnormal abdomen" sport of *Drosophila*. If cultured in a dry medium, this sport was indistinguishable from the normal form. However, if cultured in a humid environment, the segmentation of the abdomen was grossly deformed. "In a culture which progressively dries up," Hogben explained, "a decreasing number of flies manifesting the character appears. The flies which emerge last when the culture is drying up are not different from the wild type, so that in crosses conducted in the usual way any numerical results may be obtained." (ibid)

The abnormal abdomen *Drosophila* example provides another important clue in constructing Hogben's path to genotype-environment interaction, acting as something of a bridge between his discussions in 1932 and in 1933. In 1932, Hogben recognized a third class of variability resulting from the combination of a particular genetic constitution with a particular environment; the abnormal abdomen example acted to verify the existence of this class of variability. A year later, in 1933, when Hogben explicated the "interdependence" of nature and nurture for his audience at the William Withering Lectures, the abnormal abdomen example joined Krafka's (1920) bar-eye example as the two cases revealing the practical limitations on Fisher's (1918) analysis of variance. With regard to the practical significance, remember that

Hogben claimed, "A balance sheet of nature and nurture has no meaning in this sense, unless we assume that the variance of a population, if affected at all, is necessarily diminished when the environment is made more homogenous." (Hogben 1933b, 399) Hogben then utilized the abnormal abdomen example as one case showing why "Such an assumption is certainly false." (Hogben 1933a, 114)

> Imagine a large laboratory with many bottles of culture media, some dry and some moist, providing food for a mixed stock of fruit-flies, a small proportion of which belong to the mutant strain with the gene for vestigial abdomen. Keeping the stock the same, we might make the environment more homogenous in one of two ways, either making all the bottles dry or all the bottles moist. If we make all the bottles dry, the mutant gene will be incapable of manifesting its presence. Variability will be diminished with respect to the difference under consideration. If we make all the bottles moist, a larger proportion of larvae with the mutant gene will hatch out as flies with the mutant deformity. That is to say there will be an increase in variability (ibid, 115).

The crucial limitation of the abnormal abdomen example, though, was that it lacked *quantitative* data concerning the phenotype. As a result, Hogben continued, "There will be even less room for misunderstanding if we examine a metrical situation concerning which we have definite experimental knowledge." (ibid) With that, Hogben introduced Krafka's (1920) bar-eye data, displaying genotype-environment interaction quantitatively.

In light of the clear impact that Krafka's empirical results had on Hogben (providing him with the "definite experimental knowledge"), it will be fruitful to pause for a moment in order to consider Krafka's research. Three different lines of bar-eye larvae (unselected bar stock, low-bar stock, and ultra-bar stock) developed on bananas at a variety of temperatures: 15°, 16°, 17.5°, 20°, 23°, 25°, 27°, 29°, 30° and 31° C. When development was complete, the flies were etherized and the eye facets were counted with the aid of a light microscope. With the calculations completed, Krafka was able to construct the following norms of reaction:

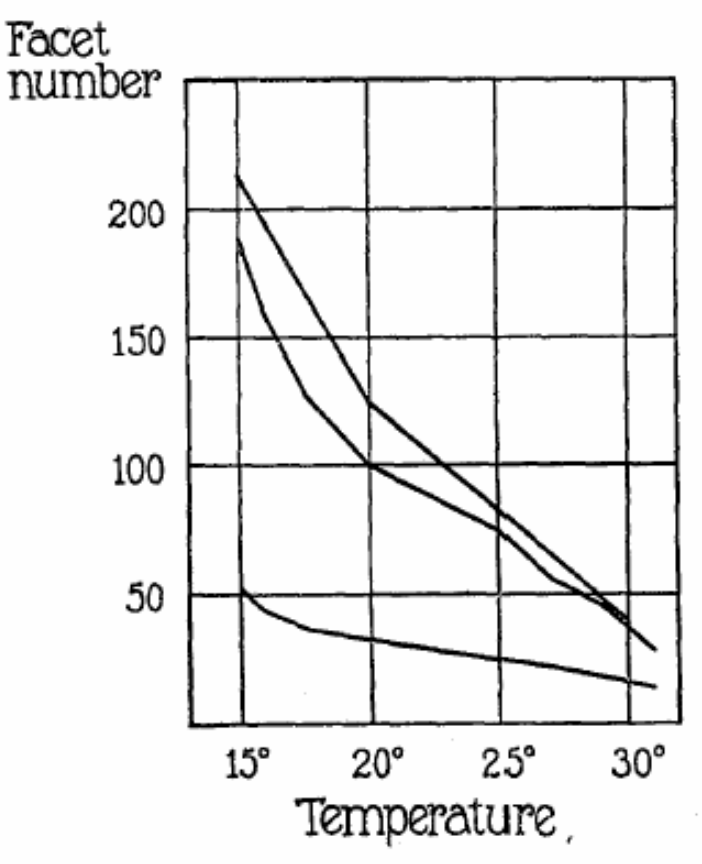

Figure 8. Krafka's norms of reaction for unselected bar stock (top), low-bar stock (middle), and ultra-bar stock (bottom) eye-facets in response to temperature. From Krafka (1920, Figure 5).[32]

Based on this graph, Krafka surmised, "We may draw two conclusions from these curves: (1) The mean facet number at any given temperature is not the same for all stocks. (2) The difference in the mean number of facets between any two temperatures is not a constant for all three stocks. In other words, the number of facets is determined by a specific germinal constitution plus a specific environment." (ibid, 419)[33] Notice that Krafka's second conclusion—

[32] Note that Hogben removed the unselected bar-stock reaction norm from his reproduction of this graph.
[33] Though Hogben's interest in Krafka's data ended with his reaction norm graphs, this data only provided the results for part 1 of the three-part essay. Krafka went on to seek "the point X in development during which the facet number-determining reaction is going on." (ibid, 434) Not surprisingly, he found different time periods at the different temperatures (beginning on the 3rd day and ending on 3 ¾ days at 27° C., and beginning on the 9th day and ending on the 11th day at 15° C.). Nevertheless, these time periods were relatively proportional to the total time of

that variation resulted from a *specific* germinal constitution plus a *specific* environment—is virtually identical to Hogben's third class of variability.

So, for Hogben, attention to individual development actually led him to recognize genotype-environment interaction as a unique source of variation in a population. He began in *Genetic Principles* by differentiating three different classes of variability: genetic, environmental, and that which "arises from the combination of a particular hereditary constitution with a particular kind of environment." (Hogben 1932a, 98) For Hogben, this last source of variation was fundamentally a developmental source of variation, resulting from differences in unique, developmental combinations of genotype and environment. In 1932, Hogben had only a qualitative example to drive this point home; the abnormal abdomen *Drosophila* strain developed quite differently in environments of different humidities in comparison to the wild type's development in these environments. When Hogben came to consider Fisher's summing of variances in 1933, though, he needed quantitative data, and he received that from Krafka (1920). Krafka's investigation of the effect of temperature on *Drosophila* development generated quantitative data revealing that "the number of facets is determined by a specific germinal constitution plus a specific environment," Hogben's third class of variation (Krafka 1920, 419). Hogben then used Krafka's data to calculate the variances for the different bar-eye stocks at the different temperatures, displaying the fact that the population variance would increase or decrease depending on which environmental temperature was chosen.

Hogben was operating in the developmental tradition of biology. The problem on which he was focused was unraveling the way in which variation in a population arose from the

---

development at these temperatures. As a result, Krafka was able to conclude, "we find that the reaction which determines the number of facets starts at the completion of 32 per cent of development and ends with the completion of 45 per cent." (ibid, 443)

relationship between genotype and environment during *individual development*. His focus was on the *causal mechanisms* of individual development. He asked, *how* do differences in genotype and differences in environment relate during individual development to generate differences in phenotype? And he employed or sought out *interventionist methodologies*, such as those undertaken by Krafka, to manipulate these variables and monitor the phenotypic outcomes. Hogben's route to genotype-environment interaction was in this developmental tradition, and the concept of interaction that he introduced bore the marks of that history. Hogben introduced what will be called the *developmental* concept of genotype-environment interaction, or $G \times E_D$. It was his "third class of variability", and it may be defined as *variation that results from differences in unique, developmental combinations of genotype and environment* (Tabery Forthcoming). The various components of the developmental tradition, now teased apart, are organized in Table 3.

Components	**Developmental Research Tradition**
Problem	Individual Development
Approach to Causation	Causal Mechanisms
Causal Question	How?
Methodology	Interventionist
Concept of Interaction	Developmental—$G \times E_D$

Table 3. The components of the developmental research tradition.

2.4. Fisher vs. Hogben: On the Importance of Genotype-Environment Interaction.

Fisher responded to Hogben's letter discussing the Krafka data.

> Dear Hogben, I think I see your point now. You are on the question of non-linear interaction of environment and heredity. The analysis of variance and covariance is only a quadratic analysis and as such only considers additive effects. Academically one could proceed in theory, though in a theory not yet

> developed, to corresponding analyses of the third and higher degrees. Practically it would be very difficult to find a case for which this would be of the least use, as exceptional types of interaction are best treated on their merits, and many become additive or so nearly so as to cause no trouble when you choose a more appropriate metric. Thus facet number shows its sweet reasonableness when measured in 'proportional units' or in other words on a logarithmic scale. However perhaps the main point is that you are under no obligation to analyse variance into parts if it does not come apart easily, and its unwillingness to do so naturally indicates that one's line of approach is not very fruitful.[34]

Fisher's appraisal of genotype-environment interaction here, along with Hogben's disregard for this appraisal in his William Withering Lectures, reveals much about their divergent views on the importance of genotype-environment interaction. Fisher understood Hogben now to be worrying about the "non-linear interaction of environment and heredity." Fisher, of course, was familiar with the problem, having taken up "Studies in Crop Variation, II" with the sole purpose of testing for such an interaction (see section 2.2.2 above, Fisher and Mackenzie 1923). With the conclusions of that study in mind, notice how Fisher responded to Hogben: Hogben's concern was written off as "academic," while "Practically it would be very difficult to find a case for which this would be of the least use, as exceptional types of interaction are best treated on their merits…" So while Fisher did acknowledge that "you are under no obligation to analyse variance into parts if it does not come apart easily," his investigation at Rothamsted led him to believe that cases where the variance did not come apart were quite "exceptional." Notice also that Fisher's response bears a striking resemblance to his review of *Genetic Principles* discussed in section 2.3.2 above, where he worried that Hogben's attention to "purely academic considerations" led to an exclusion of "aspects of more practical importance." (Fisher 1932, 147) Fisher explained that the exceptional cases of genotype-environment interaction, such as Hogben's example,

[34] Fisher to Hogben, 25 February, 1933, R. A. Fisher Papers (Series I, Hogben, L.), University of Adelaide. Quoted with the permission of the University of Adelaide Library.

could be eliminated by choosing a different scale for measuring the variables (such as a logarithmic scale for the Krafka data).

However, Fisher's *pre*-William Withering congeniality towards Hogben in this correspondence can be contrasted with a letter he wrote to J. A. Fraser Roberts two years later, after Hogben placed so much emphasis on genotype-environment interactions in his publications. While Fisher was willing to acknowledge to Hogben the limits of analyzing variance into parts "if it does not come apart easily," he wrote to Fraser Roberts on January 18th, 1935,

> …There is one point in which Hogben and his associates are riding for a fall, and that is in making a great song about the possible, but unproved, importance of non-linear interactions between hereditary and environmental factors. J.B.S. Haldane seems tempted to join in this. What they do not see is that we ordinarily count as genetic only such part of the genetic effect as may be included in a linear formula and that we make a present to the environmentalists of such variation due to the combined action of genetic and environmental causes as is not expressible in such a formula. Consequently, the more important non-linear interactions were, the more thoroughly would we underestimate the importance of genetic factors. This is, of course, another point in favour of speaking of the residue as non-genetic, rather than as environmental, though I have no doubt that in this residue the direct environmental effects are probably larger than the portion due to interaction.[35]

Fisher, here, surmised the weight he placed on genotype-environment interactions much more explicitly: they were of "possible, but unproved, importance." "Possible" because, as Fisher recognized in "Studies in Crop Variation, II," the non-linear interactions would complicate the summation of variances. But also "unproved" because in "Studies in Crop Variation, II" Fisher found no such non-linear interactions. The matter of significance was an open empirical question, and Fisher placed the burden of proof on the "environmentalists" seeking such non-linear interactions.

[35] Fisher to Roberts, 18 January 1935, quoted in Bennett (1983, 260). Quoted with the permission of the University of Adelaide Library.

Hogben, not surprisingly, came to quite a different conclusion. Krafka's (1920) research was a clear example, and Hogben took full advantage of its implications in the William Withering Lectures and the subsequent publications. Moreover, Hogben felt comfortable claiming that the reaction norm graphs revealing such genotype-environment interaction were standard; remember that he followed the discussion of the bar-eye data with the line, "The literature of experimental physiology is not wanting in examples of such divergent curves representing the measurement of a character and the strength of the environment." (Hogben 1933a, 97) As evidence, Hogben (1933b) offered the work of Taylor (1931) and Winton (1927).

With limited empirical evidence, Hogben and Fisher were free to attach quite distinct levels of importance to the empirical evidence then accumulated, leading to quite distinct conclusions concerning the importance of genotype-environment interaction. Assessing the importance of limited empirical evidence can involve any number of motivations; and, as a result, Hogben and Fisher's distinct conclusions cannot be pinned to any one motivation. Politically, Hogben's socialism naturally inclined him to favor empirical evidence supporting arguments that might justify the equalization of the environment; while Fisher's disdain for the "communists and fellow-travelers," who attacked eugenics, encouraged him to be warier of such evidence and arguments (quoted in Mazumdar 1992, 211). Turning to eugenics more directly, Hogben and Fisher's opposing perspectives on the British eugenics movement also was a potential factor affecting their respective judgments. Hogben, not surprisingly, welcomed empirical evidence complicating the statistical methodologies of eugenicists, while Fisher, not surprisingly, was critical of such evidence…especially since the methodologies were his own.

But the historical survey of Hogben and Fisher's distinct paths to genotype-environment interaction, traced out in sections 2.2 and 2.3, also revealed an *epistemological* motivation in

play, pertaining to how the concept of genotype-environment interaction figured into their respective research traditions. Fisher introduced a number of statistical innovations while at Rothamsted as part of his persistent attempts to develop methods for assessing the relative importance of heredity and the environment, the main problem of the biometric tradition. This was no isolated endeavor. Mazumdar, in investigating the *reformed* nature of Fisher's eugenics, examined the debate between Fisher and the more mainstream eugenicist, Ernest Lidbetter (Mazumdar 1992, 124-145). While Lidbetter was content to confine the Eugenics Society's Research Committee to evaluating human pedigree data on pauperism in order to emphasize the mere familial nature of the affliction, Fisher repeatedly attempted to move the Research Committee towards developing and implementing statistical techniques that could answer the question: "to what extent is the causation of pauperism to be ascribed to a) heredity b) environment?" (quoted in Mazumdar 1992, 128) Fisher, focused on the biometric "relative importance" problem, did not take genotype-environment interactions as something to be sought and studied, as if they were something of intrinsic interest. $G \times E_B$ created a potential complication for assessing the relative importance of heredity and the environment, and so it was to be considered and then either dismissed, as was the case in "Studies in Crop Variation, II," or eliminated with a transformation of scale.

Hogben, in contrast, took genotype-environment interactions to be the product of his third class of variability: that resulting from the combination of a particular genetic constitution with a particular kind of environment during the process of development. This third class of variability was, for Hogben, essentially developmental in nature and to be investigated with the tools of the developmental tradition—experimental embryology (as was the case with both the "abnormal abdomen" and bar-eye *Drosophila* studies). Just as experimental embryology was a distinct

discipline with its own inherently important results, so too was genotype-environment interaction inherently important. And as experimental embryology continued to grow, Hogben predicted, so too would empirical examples of $G \times E_D$.

The epistemological divide between Fisher and Hogben's concepts can be seen most clearly when their separate research traditions are placed side-by-side, as can be found in Table 4:

Components	**Biometric Tradition**	**Developmental Tradition**
Problem	Individual Differences	Individual Development
Approach to Causation	Causes of Variation	Causal Mechanisms
Causal Question	How Much?	How?
Methodology	Statistical	Interventionist
Concept of Interaction	Biometric—$G \times E_B$	Developmental—$G \times E_D$

Table 4. The components of the biometric and developmental research traditions.

Whatever the opposing motivations (political, social, and/or epistemological), the exchange between Fisher and Hogben evidently took its toll on their relationship. In 1932, when reviewing Hogben's *Genetic Principles*, Fisher welcomed Hogben's appointment to the Chair of Social Biology at the LSE. But in an unpublished draft of a review of Hogben's *Nature and Nurture*, Fisher began,

> Many of those, who had hopes that the establishment of a Chair of Social Biology at the London School of Economics would lead to a scientific and unbaissed [sic] attack on the social problems in this field, must by now be realising, in various degrees, their disappointment. For the functions of an advocate and of an investigator seem to be incompatible; and though one may be always amused and sometimes stimulated to thought when a brilliant journalist, such as Mr. G. K. Chesterton, sets out to show what a good forensic

> case can be made in opposition to the weight of scientific evidence and opinion, Professor Hogben lacks the charm of style needed to make confusion of thought seem luminous, or his facetiousness seem penetrating.[36]

Fisher's disdain for Hogben was by no means confined to the years of their debate. Almost 30 years later, when there was some confusion over whether an article in *Nature* was written by A. W. F. Edwards (Fisher's student) or his brother John Edwards (Hogben's student), Fisher wrote of the matter to his former colleague R. R. Race, "It was the thought that it was he [i.e., A. W. F. Edwards] that annoyed me, for the estimates published in *Nature* were manifestly incompetent, and I feared that one of my own pupils was running amok, and adding unnecessarily to darkness and confusion. However, I understand he [i.e., John Edwards] is only one of Hogben's, so all is explained."[37] A. W. F. Edwards, in fact, personally witnessed Fisher's disdain for Hogben upon the arrival of the paperback edition of Fisher's *The Genetical Theory of Natural Selection* (1958), recalling, "I was standing in the departmental office when Fisher opened the parcel of author's copies. 'Hmph,' he said at his first sight of the cover, 'Looks like a book by Hogben.'" (Edwards 1990, 278)[38]

Hogben lost no less love. In discussing the downfall of the Nazi Party in an unpublished portion of his autobiography, Hogben judged, "After the war, the Nuremberg justices of the peace had Rosenberg hanged. If I believed in hanging people for their opinions, the only extenuating circumstances I might enter with a clear conscience as a plan for mercy on behalf of

[36] R. A. Fisher Papers (Series I, Hogben, L.), University of Adelaide. Quoted with the permission of the University of Adelaide Library.
[37] Fisher to Race, 27 September 1960, R. A. Fisher Papers (Series I, Race, R.R.), University of Adelaide. Quoted with the permission of the University of Adelaide Library.
[38] I am indebted to Margaret Morrison for brining this anecdote to my attention.

the late Sir R. A. Fisher would be that he did not occupy a government post with responsibility for implementing his convictions."[39] (*LBL*, 213) And this from an avowed pacifist.

2.5. The Legacies of Fisher and Hogben: $G \times E_B$ vs. $G \times E_D$.

If the Fisher-Hogben debate had been an isolated event, then it would have been interesting in its own right, although that would have been about it. But it was not an isolated event. And, as a result, there is more than just an interesting story here. The separate concepts of interaction have had distinct legacies of their own in their separate research traditions, and the competing conceptions have faced off on numerous (sometimes acrimonious) occasions.

Hogben's $G \times E_D$ was carried into the mid-twentieth century most clearly in the work of British developmental geneticist Conrad Hal Waddington. This can be seen most clearly in Waddington's *The Strategy of the Gene* (Waddington, 1957). Waddington wanted to explain to his readers what geneticists meant by genetic and environmental influences on the phenotype. To do so, he introduced Hogben's discussion of the Krafka data and, in fact, block-quoted two full paragraphs along with the reaction norm graph from Hogben's *Nature and Nurture* where Hogben discussed the case. Reinforcing the *developmental* nature of the phenomenon, Waddington summed up, "Such a difference of degree in environmental sensitivity to the development controlled by two genotypes is spoken of as 'genotype-environment interaction'." (Waddington, 1957, 94) Like Hogben, Waddington emphasized both the importance of this phenomenon along with the mishandling of it by statistical tabulations of variance, arguing, "…after nearly half a century's development the statistical theory still has to leave out of account the contribution of genotype-environment interactions." And, "Now from the point of view of the theory of evolution such special interactions between genotypes and environments are

[39] Quoted with the permission of Leslie Hogben.

obviously by no means negligible. In fact, the whole of adaptive radiation, including the formation of local races, turns on the way in which particular genotypes fit into certain environments; that is to say, on this very factor of genotype-environment interaction." (Waddington, 1957, 100)

Waddington's emphases on the importance of a developmentally-conceived G×E, however, may be contrasted with the disregard for the concept found in the work of American population geneticist, Jay Lush, who, working in the biometric tradition, instead adopted Fisher's $G \times E_B$. In his seminal *Animal Breeding Plans* (1937), Lush brushed aside the importance of G×E in a manner reminiscent of Fisher. "It seems likely," Lush counseled, "that in general the nonadditive combination effects of heredity and environment are small in amount* and that many of those which do occur can be reduced to a negligible remainder by choosing a scale of measurements…which will show the effects of hereditary and environmental on that characteristic in their most nearly additive form." (Lush, 1937, 64) The "*" in Lush's statement directed his readers to a footnote at the bottom of the page where he continued, "For some extreme examples of nonaddtive combination effects of heredity and environment consult chapter 5 of Hogben's *Nature and Nurture*." In contrast to Waddington, then, who introduced Hogben's work as exemplifying what geneticists meant by genetic and environmental influences, Lush relegated Hogben to a footnote, as Hogben offered only "extreme examples," and, like Fisher, simply encouraged a transformation of scale to make the nuisance disappear. Thus, the competing concepts of genotype-environment interaction played out in the separate biometric and developmental traditions even after Fisher and Hogben were no longer the primary participants in the debate.

And the disputes over genotype-environment interaction did not end with Waddington and Lush. In the (in)famous IQ controversy of the 1970's, the debate focused on heritability estimates of IQ and the purported genetic basis of the difference between IQ scores for black and white populations. Critics of this genetic thesis, such as Richard Lewontin (1974) and David Layzer (1972), drew heavily on genotype-environment interaction to fundamentally undermine these heritability estimates. Employing the developmental interpretation, Layzer attacked the very meaningfulness of heritability estimates, arguing, "The information-processing skills assessed by mental tests result from developmental processes in which genetic and nongenetic factors interact continuously. The more relevant a given task is to an individual's specific environmental challenges, the more important are the effects of this interaction." (ibid, 281)

In stark contrast, Arthur Jensen, the initiator[40] of the IQ controversy and the target of Layzer's attack, described any discussions of genotype-environment interaction which invoked development as fundamentally confused. "This position," Jensen countered, "has arisen from a failure to understand the real meaning of the term 'interaction' as it is used in population genetics; but even more it is the result of a failure to distinguish between (a) the *development* of the individual organism, on the one hand, and (b) *differences* among individuals in the population." (Jensen 1973, 49, emphases in original) Jensen, like Fisher and like Lush, employed the biometric interpretation and wrote off genotype-environment interaction as exceptional since the standard biometrical methodologies did not find significant interactions.

[40] The IQ controversy is generally regarded as beginning with the publication of Jensen's "How Much Can We Boost IQ and Scholastic Achievement?" (1969).

2.6. Conclusion

This chapter examined the origins of the concept of genotype-environment interaction. In considering the origins of this concept, it was found that R. A. Fisher and Lancelot Hogben actually came to consider the concept by quite distinct routes. In developing methods for assessing the relative importance of heredity and the environment as part of the biometric tradition, Fisher came to recognize the possible complications raised by the "non-linear interaction of environment and heredity" for the summing of variances, introducing the *biometric* concept of genotype-environment interaction, or $G \times E_B$. Hogben, meanwhile, in 1932, began by considering different sources of variability in a population—a standard problem for the developmental tradition of biology. In doing so, he recognized a third class of variability (distinct from genetic or environmental variability) that resulted from the combination of a particular genetic constitution with a particular kind of environment. This source of variation was responsible for cases of genotype-environment interactions and was, for Hogben, a result of development, thus introducing the *developmental* concept of genotype-environment interaction, or $G \times E_D$.

Fisher and Hogben's distinct routes to genotype-environment interaction also led to distinct conclusions when it came to considering the importance of genotype-environment interaction. Dedicated to developing methods for assessing the relative importance of heredity and the environment, Fisher took genotype-environment interactions merely to be potential (but unproved) complications for his epidemiological techniques. Hogben, meanwhile, understood genotype-environment interactions to be of much more importance. Genotype-environment interactions were a feature of development and, as such, were to be expected in nature despite the fact that experimental embryologists were only beginning to discover them. Finally, these

separate concepts of genotype-environment interaction were traced beyond the work of Fisher and Hogben. The legacy of Fisher's $G \times E_B$ was traced through the biometric tradition in the work of Jay Lush and Arthur Jensen. And the legacy of Hogben's $G \times E_D$ was traced through the developmental tradition in the work of Conrad Hal Waddington, Richard Lewontin, and David Layzer.

3. GENOTYPE-ENVIRONMENT INTERACTION IN THE IQ CONTROVERSY

Abstract. In 1969, Arthur Jensen ignited the highly polarized IQ controversy with his appeal to genetic factors as an explanation of the difference between average IQ scores for black and white populations. One of the issues at the center of this controversy was the concept of genotype-environment interaction (G×E). Jensen and his supporters, such as Richard Herrnstein, dismissed G×E and the complications that followed from the concept pertaining to the assessment of group differences. But critics such as Richard Lewontin and David Layzer utilized G×E to buttress many of their arguments attacking Jensen's genetic hypothesis. Science studies scholars examining the IQ controversy have only perpetuated the debate surrounding G×E by simply adopting either Jensen and Herrnstein's arguments *against* or Lewontin and Layzer's arguments *for* the importance of G×E in order to argue against the opposite position. With this approach, important historical and philosophical questions have inevitably been left lingering: Why did disparate assessments of G×E exist in the IQ controversy in the first place? What was the logic of these disparate assessments? And finally, how do these disparate assessments found in the IQ controversy fit into the broader history of G×E? This chapter is an attempt at answering these lingering questions with a new conceptual framework for discussing G×E in the IQ controversy. I will argue that Jensen and Herrnstein, and Lewontin and Layzer were actually operating in different research traditions and utilizing different concepts of G×E: Jensen and Herrnstein, operating in the biometric tradition, used the *biometric* concept, or $G\times E_B$. Lewontin and Layzer, operating in the developmental tradition, used the *developmental* concept, or $G\times E_D$. The distinction between $G\times E_B$ and $G\times E_D$ provides a conceptual framework for understanding why G×E was so hotly debated in the IQ controversy, why it was hotly debated long before the IQ controversy, and why it continues to be hotly debated by historians, philosophers, and sociologists reflecting on the IQ controversy.

3.1. Introduction

The IQ controversy began in 1969 with the publication of educational psychologist Arthur Jensen's "How Much Can We Boost IQ and Scholastic Achievement?". Jensen claimed that, at least in white populations, analyses of variance showed that IQ had a relatively high heritability, implying that individual differences in IQ were largely genetically determined. Since blacks, on average, performed poorer on IQ tests, Jensen suggested that the racial gap for IQ could itself be explained genetically. Thus, egalitarian attempts to create an environment in which the racial gap for IQ disappeared were misguided; the gap was a genetic one and so would not disappear with

environmental intervention (Jensen 1969).[41] Jensen's thesis found favor with many prominent psychologists such as Richard Herrnstein, who popularized and extended Jensen's assessment of group differences (Herrnstein 1971).

Unleashed at the height of the US civil rights movement, Jensen's arguments created a furor in academia, the media, and the general public. From the academic circle, perhaps no tandem was more persistent and influential in their assault on Jensen than Harvard colleagues Richard Lewontin and David Layzer. Lewontin and Layzer drew on the concept of genotype-environment interaction (G×E) in an attempt to refute Jensen's assessment. Even if a high heritability for IQ was granted for a particular population in a particular environment at a particular time, Lewontin and Layzer reminded Jensen, this did not imply that such a high heritability would be found in a different population, or in a new environment, or at a different time. So even though whites, on average, out-performed blacks on IQ tests in the existing environment, a new environment might be encountered or created that would facilitate equal performance by both populations or that might even permit blacks to, on average, out-perform whites on IQ tests. Heritability estimates derived from analyses of variance, necessarily limited to descriptions of a particular population in a particular environment at a particular time, could not generalize beyond a limited locale; thus, Lewontin and Layzer concluded, the statistical measurement was useless. As will be shown below, Lewontin and Layzer were operating in the developmental tradition of biology. The *real* focus for geneticists, they countered in the typical developmental tradition fashion, should be the causal mechanics of the developmental genotype-environment-phenotype relationship (Feldman and Lewontin 1975; Layzer 1972a, 1972b, 1974; Lewontin 1970a, 1970b, 1974, 1975).

[41] I will only outline the arguments in this Introduction; a more detailed explication of the arguments along with the specific page citations will be provided below.

Jensen and Herrnstein, though, were unimpressed by the prospect of G×E. As will be shown below, Jensen and Herrnstein were operating in the biometric tradition of biology. In the typical biometric tradition fashion, they claimed that the analyses of variance that generated the high heritability of IQ in white populations would also reveal the presence of any G×E. However, such analyses did not reveal any interaction, so there was no use dwelling on G×E or the implications of the concept that could complicate their assessment. Heritability estimates, which focused on individual differences in a population and not on aspects of individual development, were thus very useful, and the implications derived from them were perfectly justified. Abstract speculation about what genotypic groups might do in *possible* environments, they argued, should not eclipse what the statistical estimates revealed to be occurring in the *actual* environment (Herrnstein 1973; Jensen 1970, 1972, 1973, 1975, 1976).

Because of the central role that G×E played in the IQ controversy, historians, philosophers, and sociologists of science who have examined this episode have naturally been drawn to the concept and its implications for the evaluation of group differences. The tendency, though, has been to adopt either Jensen and Herrnstein's argument *against* or Lewontin and Layzer's argument *for* the importance of G×E, and then take that argument as the final word on the matter of G×E. Two examples will suffice to display the tendency for this introduction. Neven Sesardic (1993), criticizing appeals to G×E, wrote, "…many scientists adopt a 'less than optimistic view of interactions [and think that] nonadditive interactions rarely account for a significant portion of variance' (Plomin et al. 1988, 228-229)…Others assert that the only evidence for G-E interactions comes from research on rats…and that '[n]othing like it has yet been found in human mental ability' (Jensen 1981, 124)…Concerning the heritability of

intelligence, R. Herrnstein states that 'the data from the twins reveal no interaction (in the technical sense) of heredity and environment' (1973, 180)." (Sesardic 1993, 407)

But contrast this evaluation with Allan Chase's utilization of Lewontin and Layzer's arguments from G×E to criticize Jensen's very understanding of such interactions:

> Professor Layzer, whose skills in mathematics and scientific logic and whose background and training in biology and genetics are certainly at least equal to those of Jensen and Herrnstein, concluded after a careful examination of the Jensen canon that, when it came to the interactions between gene and environment from which the phenotypes are developed, Jensen's 'remarks clearly demonstrate that he understands neither the mathematical nor the practical problems involved in the estimation of interaction effects.' A conclusion with which Dr. Layzer's colleague Richard Lewontin, the Harvard population geneticist, concurs (Chase 1980, 491).

This philosophical and historical attention to G×E is of some value, for it points to the heart of the IQ controversy. At the same time, though, simply adopting the arguments of either Jensen and Herrnstein, on the one hand, or Lewontin and Layzer, on the other hand, to criticize the opposing side leaves important historical and philosophical questions lingering: Why did disparate assessments of G×E exist in the IQ controversy in the first place? Was one side simply confused about the matter of G×E? If not, then what was the logic of the disparate assessments of G×E? In other words, how did the various components of their assessments of G×E combine to generate such different conclusions about the importance of G×E for the heritability of IQ? And finally, how did these disparate assessments found in the IQ controversy fit into the broader history of G×E? Was the debate over G×E unique to the IQ controversy or just one instantiation of a common dispute that has existed as long as the concept itself has existed? And if the latter, how do we explain the persistence of this debate?

This chapter is an attempt at answering these lingering questions with a new conceptual framework for discussing G×E in the IQ controversy. I will argue that Jensen and Herrnstein,

and Lewontin and Layzer actually employed two distinct concepts of G×E. Jensen and Herrnstein, operating in the biometric tradition, utilized the *biometric* concept of G×E, or $G \times E_B$. Lewontin and Layzer, operating in the developmental tradition, utilized the *developmental* concept of G×E, or $G \times E_D$. Operating in these different research traditions and armed with these different concepts, Jensen and Herrnstein, and Lewontin and Layzer came to quite different conceptualizations of G×E and, thus, quite different conclusions about the importance of G×E with respect to the heritability of IQ. Recognizing the distinction between $G \times E_B$ and $G \times E_D$ has a number of useful implications. First, the existence of the disparate assessments found in the IQ controversy can be explained with reference to the disparate concepts and different research traditions of Jensen, Herrnstein, Lewontin, and Layzer. Moreover, attempting to dismiss either side of the debate simply by writing off those individuals as confused about G×E overlooks the fact that both sides drew on the existent empirical data on G×E and provided multiple reasons for their respective conclusions about the concept's significance. And finally, the conflict between the different interpretations of G×E found in the IQ controversy can be recognized as nothing new for G×E; in fact, the distinct concepts existed and, indeed, competed long before the question of the heritability of IQ arose in the 1970s, as was shown in the last chapter. Thus the debate concerning G×E found in the IQ controversy can be understood as just one instantiation of this persistent trend, and this persistence can itself be explained with reference to the existence of both $G \times E_B$ and $G \times E_D$ throughout the history of G×E.

In the next section I will outline the debate between Jensen, Herrnstein, Lewontin, and Layzer. In this descriptive portion of the chapter, I will lay out their arguments and also detail the place of G×E in each of those arguments. Section 3.3 is then devoted to a closer examination of how science studies scholars have evaluated this debate. Examples were given above, but with a

more detailed picture of Jensen, Herrnstein, Lewontin, and Layzer's arguments in place from section 3.2, a more detailed picture can be given of how these science studies scholars have focused on G×E simply by adopting the arguments provided by the scientists. This approach will be contrasted with the one I take up in section 3.4, where the distinction between $G\times E_B$ and $G\times E_D$ will be introduced and used to evaluate the debate between Jensen, Herrnstein, Lewontin, and Layzer. In addition to specifying the character of each concept here in its respective research tradition, I will show how employing either $G\times E_B$ or $G\times E_D$ led Jensen and Herrnstein, and Lewontin and Layzer to come to different conclusions concerning common questions about how G×E should be conceptualized and about how important the concept was for the heritability of IQ.

3.2. Genotype-Environment Interaction in the IQ Controversy

"Compensatory education has been tried and it apparently has failed….The chief goal of compensatory education—to remedy the educational lag of disadvantaged children and thereby narrow the achievement gap between 'minority' and 'majority' pupils—has been utterly unrealized in any of the large compensatory education programs that have been evaluated so far." (Jensen 1969, 2-3) Thus began Jensen's appraisal of the egalitarian attempts to eliminate the "achievement gap" between advantaged, white children and disadvantaged, black children. The diagnosed failure, Jensen continued, forced one to consider the question: Why has compensatory education failed?

Jensen's answer to this question came from evaluating "the nature of intelligence" (ibid, 5). Intelligence (or, more specifically, Charles Spearman's (1904) "general intelligence"—*g*) was normally distributed in human populations. This meant that, like physical traits such as height, a

large sample of individuals tested for intelligence with a measure such as IQ would be distributed across the possible scores with the largest number congregating about the mean score (100, in the case of IQ) and gradually decreasing as scores deviated more and more from this mean in either direction. The result was a bell curve (ibid, 20-28). The fact that IQ scores were distributed in such a manner, Jensen argued, revealed that the trait being measured with such scores—intelligence—was polygenic, meaning that individual differences in the trait were the result of multiple genes, whose effects were small, similar, and cumulative (ibid, 33). The distribution also allowed for the statistical quantification of this variation by means of the concept of *variance*. Following in the biometric tradition introduced in the last chapter, Jensen explained that variance was an index of the total amount of variation among scores, and since variance accounted for variation on an additive scale, the total variance of a distribution of scores could be partitioned into a number of separate components, each of these components being due to a factor contributing a specifiable proportion of the variance, and all the variance components adding up to the total variance. Jensen was quick to pay his debt to one of the fathers of the biometric tradition in which he was operating: "The mathematical technique for doing this, called 'the analysis of variance', was invented by Sir Ronald Fisher, the British geneticist and statistician. It is one of the great achievements in the development of statistical methodology." (ibid, 28)

The additive nature of the analysis of variance allowed one to treat the total variance as the sum of a number of separate variance components. With these separate variance components considered together, the result was the following equation:

$$V_P = (V_G + V_{AM}) + V_D + V_i + V_E + 2Cov_{HE} + V_I + V_e \quad (3.1)$$

V_P referred to the total phenotypic variance of the trait in the population; V_G captured the variance due to gene effects which were additive; V_{AM}, the variance due to assortative mating, was conjoined with V_G since it directly affected the proportion of this other component; V_D referred to the variance due to dominance deviation; and V_i was the component of variance attributable to the interaction of genes (epistasis). Meanwhile, V_E accounted for the environmental variance; $2Cov_{HE}$ represented the variance due to genotype-environment correlation; and V_I was the variance due to statistical interaction of genetic and environmental factors (G×E). Finally, V_e was included to capture variance due to unreliable errors of measurement (ibid, 34).

For Jensen's purposes, though, this equation was only the beginning. He was interested in the inheritance of intelligence and, thus, the proportion of the variation of this trait in the population due to variation in genotypes. So another concept from the biometrical tradition was employed—*heritability* (h^2). This concept was defined as the proportion of the total phenotypic variation arising from variation due to heredity:[42]

$$h^2 = [(V_G + V_{AM}) + V_D + V_i]/V_P - V_e \quad (3.2)$$

Heritability was the crucial concept and statistical measure for Jensen's assessment of the causes of variation in intelligence. Based largely on the research of Sir Cyril Burt (Burt 1955, 1958, 1966), "a 'must' for students of individual differences" according to Jensen (ibid, 33), Jensen placed the estimates of heritability for intelligence in white populations at roughly 0.8 (ibid, 46-

[42] It is important to note here that this is the definition of heritability in the *broad* sense. Following Jay Lush's distinction, this concept can be contrasted with heritability in the *narrow* sense, which only accounts for the proportion of total phenotypic variance arising from the additive genetic component (V_G) (Lush 1943). Heritability in the broad sense is often referred to as the measure of genetic determination, and it is the concept that was at the heart of the debates in the IQ controversy. Heritability in the narrow sense is instead a measure of how successful selection will be for the particular trait under investigation and is of more interest to animal breeders than to human geneticists and psychologists. As a result, it played a less pivotal role in the IQ controversy.

59). This high heritability, Jensen continued, revealed a significant role for genetic factors in explaining the individual differences in intelligence.

Unfortunately, according to Jensen, "…the possible importance of genetic factors in racial behavioral differences has been greatly ignored, almost to the point of being a tabooed subject, just as were the topics of venereal disease and birth control a generation or so ago." (ibid, 80) Jensen's goal was to break down this taboo. On the average, blacks tested about one standard deviation (15 IQ points) below the average of the white population in IQ. Moreover, even when socioeconomic level was controlled (obviously of relevance when blacks are disproportionately represented among lower socioeconomic levels), the average difference only reduced to about 11 IQ points (ibid, 81). "There is an increasing realization among students of the psychology of the disadvantaged," Jensen surmised, "that the discrepancy in their average performance cannot be completely or directly attributed to discrimination or inequalities in education. It seems not unreasonable, in view of the fact that intelligence variation has a large genetic component, to hypothesize that genetic factors may play a part in this picture." (ibid, 82)

Jensen's controversial genetic hypothesis found favor with other scientists such as the Harvard psychologist Richard Herrnstein.[43] In 1971, Herrnstein distilled Jensen's argument for a more popular audience in the pages of the *Atlantic Monthly*, in an article entitled "I.Q." (Herrnstein 1971). After historically introducing his readers to the research of Francis Galton on inheritance and Alfred Binet and Theodore Simon on mental testing, Herrnstein turned to the "cautious and detailed" analysis to be found in Jensen's 1969 (ibid, 55). Far from being "extreme in position or tone," Herrnstein claimed Jensen's article simply summarized what was already widely recognized in scientific communities. "Not only its facts but even most of its conclusions

[43] Other prominent advocates for Jensen's argument were the Nobel laureate William Shockley and the British psychologist Hans J. Eysenck (Eysenck 1971, 1973; Eysenck and Fulker 1979; Shockley 1972, 1978).

are familiar to experts," Herrnstein wrote; "Jensen echoes most experts on the subject of the I.Q. by concluding that substantially more can be ascribed to inheritance than environment." (ibid)

Herrnstein's own contribution to the debate was to go beyond the question of *racial* differences and discuss the question of *class* differences more generally. Herrnstein, to make this extension, offered a syllogism: If differences in mental abilities are inherited, and if success requires those abilities, and if earnings and prestige depend on success, then social standing (which reflects earnings and prestige) will be based to some extent on inherited differences among people (ibid, 58, 63). This conclusion, Herrnstein continued, had important implications for the future of class structure. Social mobility allowed individuals with superior innate capacities to climb up the class hierarchy, but individuals with inferior innate capacities could not make such strides; over time, the upper classes would consist entirely of the intellectually superior individuals, while the lower classes would, in turn, consist entirely of the intellectually inferior individuals. And because intelligence is highly heritable, the associated traits relevant to class status (success, earnings, prestige) would be highly heritable too, thus locking individuals from the higher and lower classes into their domains by virtue of their genetically-governed intelligence. "What is most troubling about this prospect," Herrnstein warned, "is that the growth of a virtually hereditary meritocracy will arise out of the successful realization of contemporary political and social goals. The more we succeed in achieving relatively unimpeded social mobility, adequate wealth, the end of drudgery, and wholesome environment, the more forcefully does the syllogism apply." (ibid, 63-64)

3.2.1. The Argument from Genotype-Environment Interaction

Critics attacked Jensen and his supporters from a variety of angles, questioning, for example, the validity of the data borrowed from Burt, the inherent biases in intelligence testing, the reality of 'race' as an actual biological entity, and the assessment of compensatory education's failure.[44] Lewontin, though, sought to undermine the very methodological foundations of Jensen's genetic hypothesis. The Harvard geneticist drew on the concept of G×E to attack the implications drawn from heritability estimates by Jensen and his supporters. Lewontin began his assault in a 1970 article for the *Bulletin of the Atomic Scientists*. While Jensen (1969) took the importance of genetic factors affecting intelligence to create a problem for attempts at environmental intervention on the trait (i.e., compensatory education), Lewontin argued that a genetic component to intelligence in no way created such a problem. "Let it be entirely genetic," Lewontin granted (Lewontin 1970a, 8). "Does this mean that compensatory education, having failed, must fail? The supposition that it must arises from a misapprehension about the fixity of genetically determined traits." (ibid) Lewontin pointed to the "abnormalities of development" to make this point:

> It was thought at one time that genetic disorders, because they were genetic, were incurable. Yet we now know that inborn errors of metabolism are indeed curable if their biochemistry is sufficiently well understood and if deficient metabolic products can be supplied exogenously. Yet in the normal range of environments, these inborn errors manifest themselves irrespective of the usual environmental variables. That is, even though no environment in the normal range has an effect on the character, there may be special environments, created in response to our knowledge of the underlying biology of a character, which are effective in altering it (ibid).

[44] A reference source on these various criticisms can be found in Aby and McNamara's (1990) *The IQ Debate: A Selective Guide to the Literature*.

Jensen claimed that an environment of abundance would do little to elevate the lower IQ scores of blacks in relation to whites because of the genetic basis of the trait indicated by the high heritability estimates; but Lewontin countered, "It is empirically wrong to argue that if the richest environment experience we can conceive does not raise I.Q. substantially, that we have exhausted the environmental possibilities." (ibid) Determining the environments available to individuals, Lewontin emphasized, was a social matter not a biological one. Thus, "In answer to Prof. Jensen's rhetorical question 'How much Can We Boost IQ and Scholastic Achievement?' I say 'As much or as little as our social values may eventually demand.'" (Lewontin 1970b, 25)

In this early critique of Jensen's assessment of intellectual differences between black and white populations Lewontin introduced several points that would arise time and again in his future discussions of the methodological foundations of heritability estimates and the implications derived from those estimates: (a) the importance of seeking the causal mechanics of development along with the need to admit ignorance when such causal mechanics were unknown, and (b) the importance of emphasizing *possible* environments as a source of potentially new phenotypic outcomes. Lewontin developed these points later in the decade when he challenged analyses of variance and heritability estimates more generally, revealing Jensen's employment of the statistical methodologies as just one instance of a more general problem in statistical biology. In his influential "The Analysis of Variance and the Analysis of Causes," Lewontin (1974) pointed to a "problem of causation" for the analysis of variance: "the problem of analyzing into separate components the interaction between environment and genotype in the determination of phenotype. Here…we recognize that all individuals owe their phenotype to the biochemical activity of their genes in a unique sequence of environments and to developmental events that may occur subsequent to, although dependent upon, the initial action of the genes."

(ibid, 401) The pseudo-question concerning "nature versus nurture," Lewontin claimed, arose because "It was supposed that the phenotype of an individual could be the result of *either* environment *or* genotype, whereas we understand the phenotype to be the result of *both*." (ibid, emphases in original)

To justify this point, Lewontin pointed to the interventionist research of Conrad Hal Waddington (1953) and J. M. Rendel (1959) on developmental canalization. This developmental canalization could be graphed with norms of reaction revealing G×E, such as those Lewontin provided in Figure 9, where the phenotype (P) was plotted as a function of the environment (E) and two different genotypes.[45] The reaction norms graph in Figure 9(e), Lewontin explained, was common for enzyme activity, where genotypes were displaced horizontally based on having different temperature optima (the environmental variable). These graphs, with their explicit G×E, had important implications for analyses of variance that only investigated sources of variation in single or limited environments because the more general genotype-environment-phenotype relationship was missed. Lewontin warned that "if the temperature distribution is largely to either side of the crossover point between these two genotypes, there will be very large components of variance for both genotype and environment and a vanishingly small interaction component; yet over the total range of environments exactly the opposite is true!" (ibid, 407) Continuing, Lewontin explained,

> Figure [9]e also shows a second important phenomenon, that of differential phenotypic sensitivity in different environmental ranges. At intermediate temperatures there is less difference between genotypes and less difference between the effect of environments than at more extreme temperatures. This is the phenomenon of canalization and is more generally visualized in figure [9]f. Over a range of intermediate phenotypes there is little effect of either genotype or environment, while outside this zone of canalization

[45] For a history of the reaction norm concept, see Sarkar (1999)

phenotype is sensitive to both. …The sensitivity of phenotype to both environment and genotype is a function of the particular range of environments and genotypes. For the programmatic purposes of human genetics, one needs to know more than the components of variation in the historical range of environments (ibid).

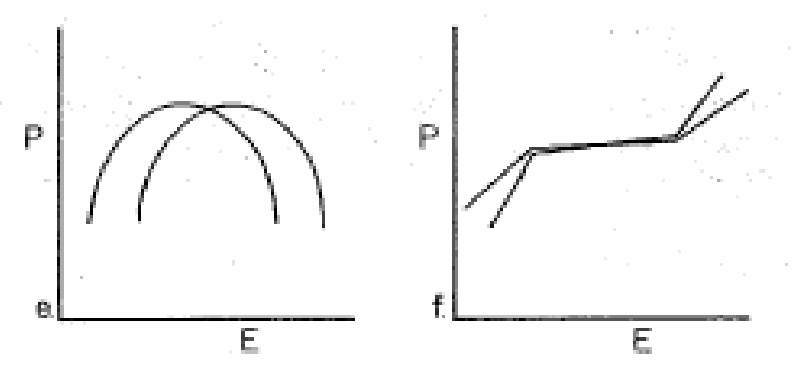

Figure 9. Lewontin's hypothetical reaction norm graphs for phenotypic traits caused by canalization. From Lewontin (1974, Figures 1e and 1f). Reprinted with the permission of the University of Chicago Press.

The complex interdependence between the genotypic and environmental distributions in Figure 9 can be contrasted with those Lewontin provided in Figure 10. As Jensen readily admitted, the analysis of variance necessarily employed an assumption of additivity between genotypic and environmental sources of variation. Such an assumption was accurate if the norms of reaction for the phenotypic trait resembled those found in Figure 10, but was such an assumption empirically justifiable? Lewontin argued it was not, and so additivity was often simply assumed "because it suits a predetermined end." (ibid, 409) This was the suspicious reasoning Lewontin tacked on Jensen; Figure 10(g) was "the hypothetical norm of reaction for IQ taken from Jensen (1969). It purports to show the relation between environmental 'richness' and IQ for different genotypes." (ibid) However, Lewontin ridiculed such a picture: "While there is not a scintilla of evidence to support such a picture, it has the convenient properties that

superior and inferior genotypes in one environment maintain that relation in all environments, and that as environment is 'enriched,' the genetic variance (and therefore the heritability) grows greater." (ibid) Lewontin concluded sarcastically, "This is meant to take care of those foolish egalitarians who think that spending money and energy on schools generally will iron out the inequalities in society." (ibid)

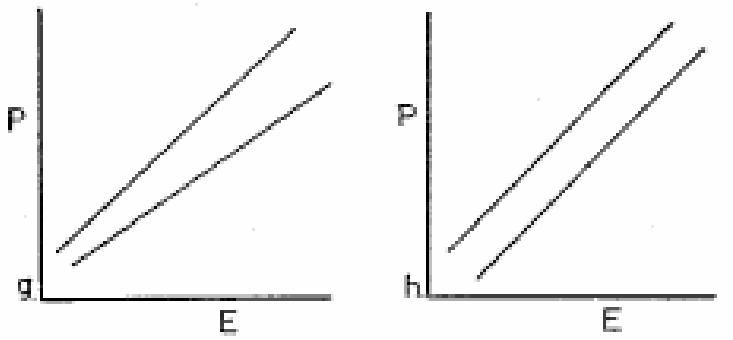

Figure 10. Lewontin's hypothetical reaction norm graphs for phenotypic traits caused by an additive genotype-environment relation. From Lewontin (1974, Figures 1g and 1h). Reprinted with the permission of the University of Chicago Press.

Lewontin reiterated the importance of the causal mechanics of development, possible environments, and G×E a year later with Marcus Feldman in their criticism of statistical genetics' "Heritability Hang-up" (Feldman and Lewontin 1975). Feldman and Lewontin admitted that it was possible to estimate variation due to G×E in controlled animal and plant breeding experiments. But they continued by emphasizing the importance of potential environments that may have been neglected in such experiments: "Nevertheless, genotype-environment interaction remains a serious problem even in agricultural applications. If varieties are tested under a particular range of conditions, or a selection program is carried out over a limited range of environments, the selected material may be totally inappropriate for other

conditions." (ibid, 1164) To reveal the thrust of this line of reasoning, Feldman and Lewontin provided the reaction norm graph in Figure 11 with extreme G×E present. "Obviously, both genotype and environment influence the phenotype in this example," the two pointed out. And again the emphasis was on the mistaken picture that would be gleaned if an analysis of variance tested only E_1:

> …if the environments are symmetrically distributed around E_1 (Fig. [3.3]), there will appear to be no average effect of genotype, while if the population is weighted toward an excess of G_1, the average phenotype across environments will be constant, as is shown by the dashed line. Thus the environmental variance depends on the genotypic distribution, and the genotypic variance depends on the environmental variance. This very important interdependence means that for a character like IQ, where the norm of reaction, the present genotypic distribution, and the present environmental distribution are not known, we cannot predict whether an environmental change will change the total variation (ibid, 1166).

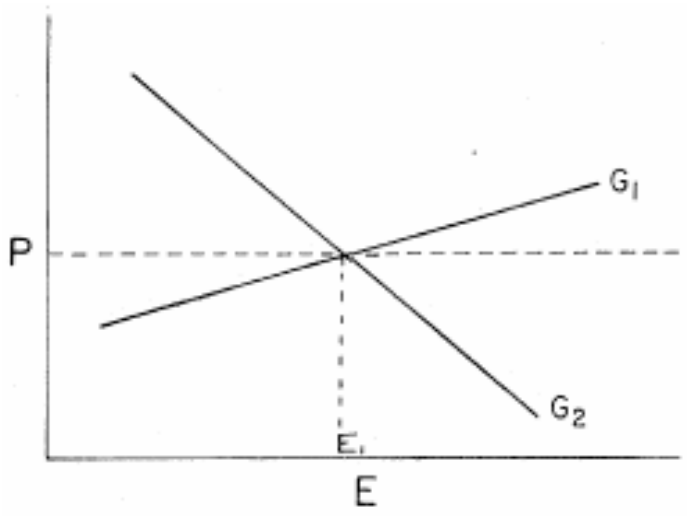

Figure 11. Lewontin and Feldman's hypothetical reaction norm graph. From Feldman and Lewontin (1975, Figure 1).

Importantly, Feldman and Lewontin charged Jensen and Herrnstein with ignorance of G×E and its implications for their heritability estimates. They chided, "This situation is ignored by both Jensen (1969) and Herrnstein (1971), whose discussion does not take account of this possible form of genotype-environment interaction." (ibid)

Lewontin drew on the concept of G×E with norm of reaction graphs to make clear the importance he placed on the need to consider the causal mechanics of developmental biology and possible, as-yet-untested environments in any discussion of variation and group differences. G×E was then used, in turn, to attack the statistical analyses of variance for their ignorance of such causal mechanics and their confinement to limited, tested environments. Because additive analyses of variance overlooked these fundamental components, Lewontin questioned whether the analyis of variance had *anything at all* to offer such research on variation, concluding his 1974, "The simple analysis of variance is useless for these purposes and indeed it has no use at all. In view of the terrible mischief that has been done by confusing the spatiotemporally local analysis of variance with the global analysis of causes, I suggest that we stop the endless search for better methods of estimating useless quantities. There are plenty of real problems." (Lewontin 1974, 410) This was a conclusion he reiterated with Feldman: "…relations between genotype, environment, and phenotype are at base mechanical questions of enzyme activity, protein synthesis, developmental movements, and paths of nerve conduction. We wish, both for the sake of understanding and prediction, to draw up the blueprints of this machinery and make tables of its operating characteristics with different inputs and in different milieus. For these problems, statistical descriptions, especially one-dimensional descriptions like heritability, can only be poor, and, worse, misleading substitutes for pictures of the machinery." (Feldman and Lewontin 1975, 1167-1168)

Lewontin's assault on the methodological foundations of heritability estimation was not universally welcomed. In fact, association with Lewontin at this time apparently had its own unique repercussions: Shortly after co-authoring the aforementioned article with Lewontin, Feldman wrote to his collaborator and explained that his recent job search was stymied because an administrator in the medical school blocked the job offer because of their article in *Science*. Apparently, Feldman learned, the administrator had programs funded for the study of the genetics of disorders such as alcoholism and schizophrenia.[46] Lewontin, though sorry for the inconvenience, seemed not at all surprised by the event: "Perhaps you will realize now that a close association with me has some real disadvantages, and that you might be wise to be a little more circumspect. I feel extremely bad about this and I urge you to consider the possibility that in the future you should be more cautious, especially where I am concerned."[47] (Feldman was ultimately offered the position after all.[48])

Lewontin's project was not entirely negative, devoted only to tearing down the heritability estimates utilized by Jensen and university administrators endorsing research on the genetics of alcoholism and schizophrenia. He also offered a positive thesis about what geneticists ought to be studying. As suggested in the quote above from Feldman and Lewontin (1975), Lewontin emphasized the importance of interventionist methodologies that revealed the *causal mechanics* of the developmental genotype-environment-phenotype relationship, which could be represented with norms of reaction: "The real object of study both for programmatic and theoretical purposes is the relation between genotype, environment, and phenotype," Lewontin wrote in 1974. "This is expressed in the *norm of reaction*, which is a table of correspondence

[46] Feldman to Lewontin, 20 November 1976, Richard Lewontin Papers (Feldman file), American Philosophical Society (APS) Library, Accession Number B L59p.
[47] Lewontin to Feldman, 9 December 1976, Lewontin Papers (Feldman file). Quoted with the permission of the American Philosophical Society.
[48] Feldman to Lewontin, 21 December 1976, Lewontin Papers (Feldman file).

between phenotype, on the one hand, and genotype-environment combinations on the other." (Lewontin 1974, 404, emphasis in original) As an example of a successful instance of such a reaction norm approach to genetics, Lewontin referred his readers to the early work of his mentor, the Russian geneticist Theodosius Dobzhansky. Dobzhansky's study of different *Drosophila* genotypes' viability developing at different temperatures provided the norms of reaction found in Figure 12, which also importantly offered for Lewontin an empirical instance of G×E (Dobzhansky and Spassky 1944).

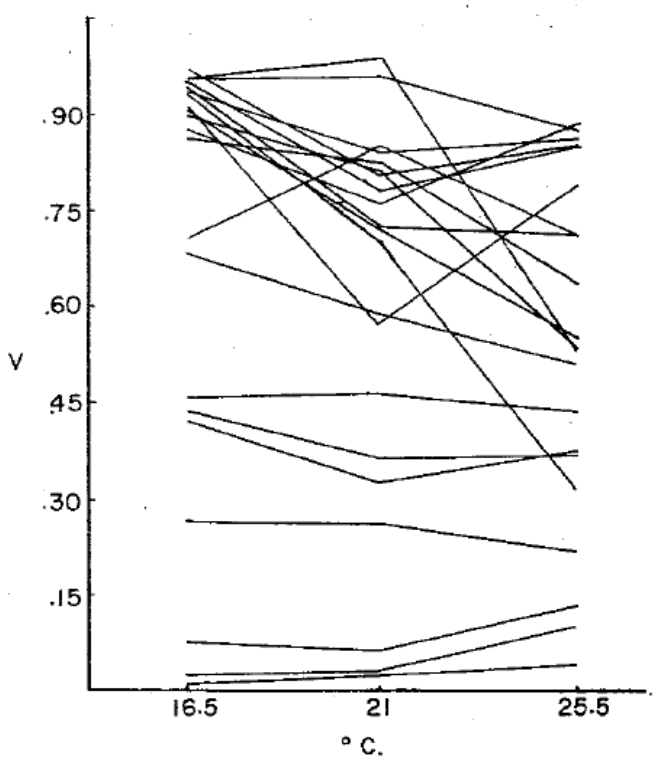

Figure 12. Lewontin's actual reaction norms for viability of fourth chromosome homozygotes of *Drosophila pseudoobscura*. From Lewontin (1974, Figure 2). Reprinted with the permission of the University of Chicago Press.

Despite the fundamental importance Lewontin placed on the causal mechanics of development, he was simultaneously concerned that perhaps he and Dobzhansky were two of the very few geneticists that properly recognized this importance at the time of the IQ controversy. Writing to Dobzhansky in May 1973 following the publication of Dobzhansky's *Genetic Diversity and Human Equality* (1973), Lewontin worried,

> Just a couple of comments on the book. First, you remain the only geneticist writing on general subjects and even one of the very few writing on technical subjects who says correct things about environment and genotype. The notion of the norm of reaction has simply failed to permeate the general textbook writings of our colleagues. As a result, they give all the wrong impression. I have recently done a survey of textbooks and find that among them only Sinnott, Dunn, and Dobzhansky [(1958)] makes a suitable presentation of this topic. I was delighted to see it carried through in your latest book. Why is it that most geneticists do not understand that the phenotype is a developmental process?[49]

Even though Lewontin argued that norms of reaction were what geneticists ought to study, he also readily admitted that such a focus was, for all practical purposes, impossible for human geneticists in the 1970's. Lewontin conceded, "In man, measurements of reaction norms for complex traits are impossible because the same genotype cannot be tested in a variety of environments." (Lewontin 1974, 409) In his review "Genetic Aspects of Intelligence," published in 1975, Lewontin delivered a similar verdict, beginning, "Indeed, this study of norms of reaction is the proper object of research—if we are interested in knowing how various historical changes in human social organization and educational practice will affect human behavior. This is the only correct sense in which we can study the 'nature-nurture' problem, the problem of the interacting genetic and environmental causes (Lewontin 1974). It is in this sense that we analyze the genetics of larval viability in *Drosophila* (Dobzhansky and Spassky (1944)." (Lewontin

[49] Lewontin to Dobzhansky, 2 May 1973, Theodosius Dobzhansky Papers (Lewontin file), American Philosophical Society (APS) Library, Accession Number B D65. Quoted with the permission of the American Philosophical Society.

1975, 387) But then Lewontin quickly continued, "But even this level of investigation is denied us for human traits, most especially behavioral traits, because we simply cannot replicate human genotypes over and over and follow their development in different environments. Indeed, we do not even know what we mean by environment in this case since it presumably includes the overwhelming complexity of social milieu and is itself an autocorrelated developmental process." (ibid)

Such difficulties as these, however, did not prevent Lewontin from passing friendly judgment on *individual* norms of reaction: On the occasion of Dobzhansky's 75th birthday, Lewontin wrote to his mentor, "It is 1975 and that means that in a few days it will be your 75th birthday. I write to send my filial and paternal love on this great occasion. If there were a God, I would thank Him for bringing you so brilliantly through three-quarters of a century. As it is, you have only your genes and your environment to thank. With such a norm of reaction I have no doubt that you will reach 100!"[50] (Unfortunately, Lewontin was incorrect in this assessment; Dobzhansky died later that same year.)

Lewontin's emphasis on G×E and its implications for heritability estimates was not confined to scholarly journals. In an article for the *Boston Phoenix* entitled "The Brains Do Battle in I.Q. Controversy," journalist Paul Wagman outlined the assault on Jensen and Herrnstein's claims concerning the wide-spread acceptance by psychologists of their genetic hypothesis, writing, "Only people who have made an intensive study of quantitative genetics, says Lewontin, are prepared to understand the subject of heritability well enough to make such estimates. Layzer, in a paper he has prepared for *Science*, maintains that the analyses which have led to the consensus cited by Herrnstein are shot through by systematic errors." (Wagman 1973,

[50] Lewontin to Dobzhansky, 7 January 1975, Dobzhansky Papers (Lewontin file), APS. Quoted with the permission of the American Philosophical Society.

28) A photograph worth 1000 words accompanied Wagman's report in which Lewontin displayed his frustration with his opponents (Figure 13). And again G×E figured prominently in Lewontin's frustration; he stood adjacent to a blackboard on which a hypothetical reaction norm graph was drawn with three genotypes. Not surprisingly, the graph displayed significant G×E: Genotype 1 was superior to both Genotype 2 and Genotype 3 in environments to the left side of Lewontin's graph, but in environments to the right side of the graph, Genotypes 2 and 3 climbed high above Genotype 1. G×E was thus the primary take-home lesson for even a journalist and photographer covering the IQ controversy.

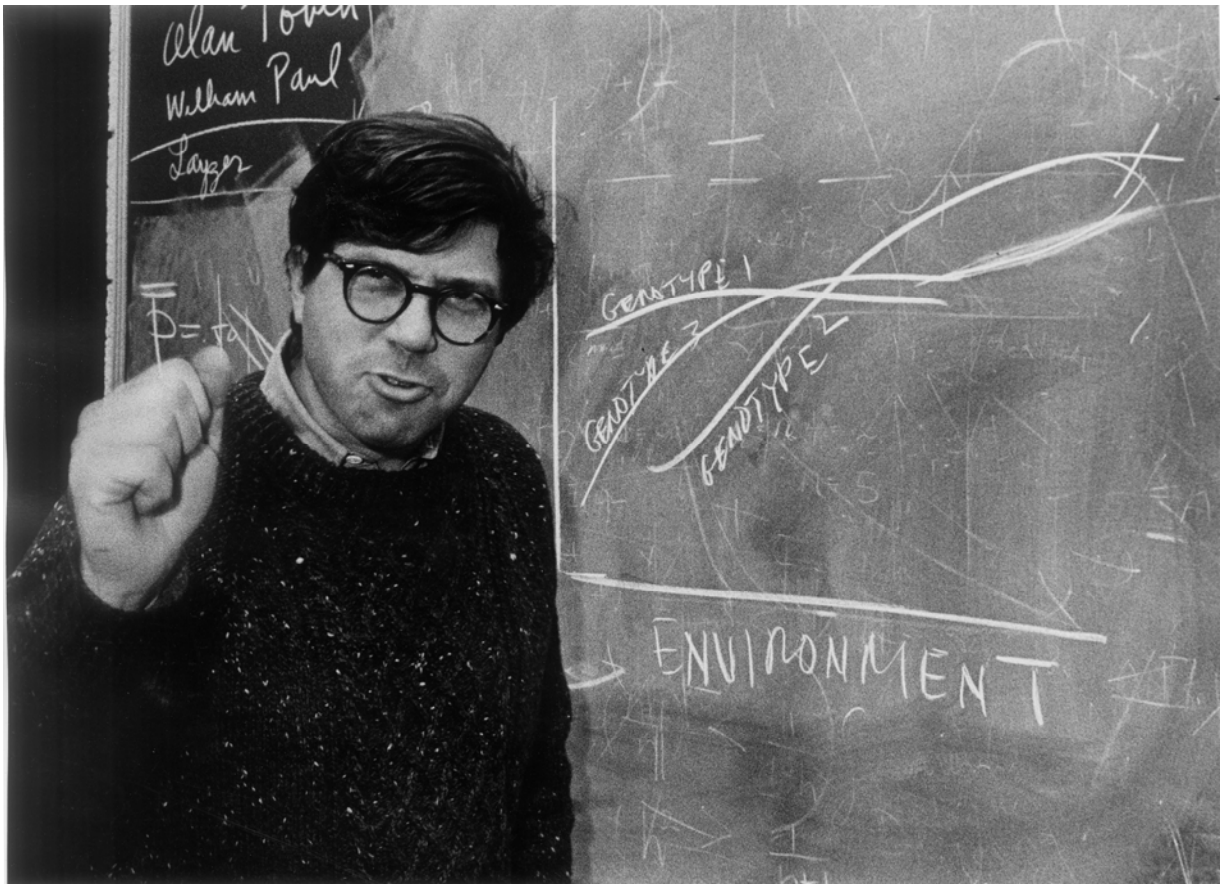

Figure 13. Photograph of Richard Lewontin from 1973 *Boston Phoenix* article "The Brains Do Battle in I.Q. Controversy." Reprinted by permission of photographer, Ken Kobre.

Behind Lewontin's head in the photograph above is Layzer's name highlighted prominently; and, in his reporting of the IQ controversy, Wagman (1973) united Lewontin and Layzer as the prominent critics of Jensen and Herrnstein. It was not surprising that Lewontin would mention his colleague at Harvard when emphasizing the importance of considering G×E in discussions of heritability estimates, for Layzer also drew on the concept of G×E to criticize Jensen in a number of articles throughout the early 1970's. Layzer first took up the matter in an exchange with Jensen on the pages of the first volume of the journal *Cognition* in 1972. Like Lewontin, Layzer drew attention to the assumed additivity between genotype and environment implemented by heritability estimates. He warned, "The assumption that genetic and environmental factors contribute additively and independently to a phenotypic character is, on general grounds, highly suspect. From a purely mathematical point of view, additivity is an exceedingly special property. Moreover, a character that happens to have this property when measured on one scale would lose it under a nonlinear transformation to a different scale of measurement. Additivity is therefore a plausible postulate only when there exists some specific biological justification for it." (Layzer 1972a, 275) For complex traits, however, Layzer doubted that such a biological justification existed: "For complex animal characters there is little reason to expect additivity and independence to prevail. On the contrary, such characters usually reflect a complicated developmental process in which genetic and environmental factors are inextricably mingled." (ibid)

The information processing skills loosely measured by IQ, Layzer went on, were an instance of such a complex animal character. "Information-processing skills, like other skills, are not innate, but develop over the course of time. What is the nature of this development?," Layzer

asked. Answering his own rhetorical question, Layzer explained that complex skills, such as skiing or playing a piano, were acquired in succession via a number of intermediate techniques, and each of these allowed one to perform competently at a certain intermediate level of difficulty before progressing to the next level. Layzer, citing the work of Jean Piaget (1952), claimed that cognitive structures developed in the same way: "Each new structure is always more highly organized and more differentiated than its predecessor. At the same time it is more adequate to a specific environmental challenge." (ibid, 280) Layzer compared this development to the building of a house; logic and the laws of physics would require that the various stages be completed in a particular order (foundation, then frame, then walls, then roof), but the skills of the builder, the available materials, the builder's intentions, and the nature of the environmental challenge would all also govern the construction. "Similarly," Layzer continued, "although cognitive development is undoubtedly strongly influenced by genetic factors, it represents an adaptation of the human organism to its environment and must therefore be strongly influenced by the nature of the environmental challenge. Thus we may expect cultural factors to play an important part in shaping all the higher cognitive skills, for the environmental challenges that are relevant to these skills are largely determined by cultural context." (ibid)

This complex interplay between genetic and environmental factors opened the door to Layzer's discussion of interaction:

> If intelligence, or at least its potentially measurable aspects, can be identified with information-processing skills and if the preceding very rough account of how these skills develop is substantially correct, then it seems highly unlikely that scores achieved on mental tests can have the mathematical properties that we have been discussing—properties needed to make 'heritability of IQ' a meaningful concept. The information-processing skills assessed by mental tests result from developmental processes in which genetic and nongenetic factors interact continuously. The more relevant a given task is to an individual's specific environmental challenges, the more important are the effects of this interaction (ibid, 281).

With this argument in mind, Layzer dubbed 'the heritability of IQ' a pseudo-concept, on par with 'the sexuality of fractions,' and the 'analyticity of the ocean.' (ibid, 294)

Wagman's report on the IQ controversy mentioned a paper Layzer had "prepared for *Science*," which revealed that the consensus Herrnstein and Jensen pointed to in support of their genetic hypothesis was "shot through with systematic errors." (Wagman 1973, 28) Layzer's "Heritability Analyses of IQ Scores: Science or Numerology?" came out a year later, the same year as Lewontin's "The Analysis of Variance and the Analysis of Causes." Layzer, here, focused on the limitations of the heritability concept and, in so doing, offered a reaction norm graph (Figure 14) reminiscent of the reaction norm graphs Lewontin drew on so heavily to deliver his own attacks on heritability estimates. Layzer began, "Genetic differences may influence the development of a trait in qualitatively distinct ways." (Layzer 1974, 1260) Referring to Figure 14, Layzer pointed out that the three genotypes had different thresholds, different slopes, and different final values. "Heritability estimates do not take such qualitative distinctions into account," Layzer charged (ibid). If the environment is quite narrow, confined to the area around y_1 for example, h^2 would be close to unity. But, Layzer argued, this conclusion only revealed the limited scope of heritability estimates, for "…in these circumstances the phenotypic variance could reasonably be considered to be largely environmental in origin since it is much greater than the phenotypic variance that would be measured in an environment ($y = y_2$) that permitted maximum development of the trait, consistent with genetic endowment. This point has been elaborated by R. C. Lewontin (1970a)."

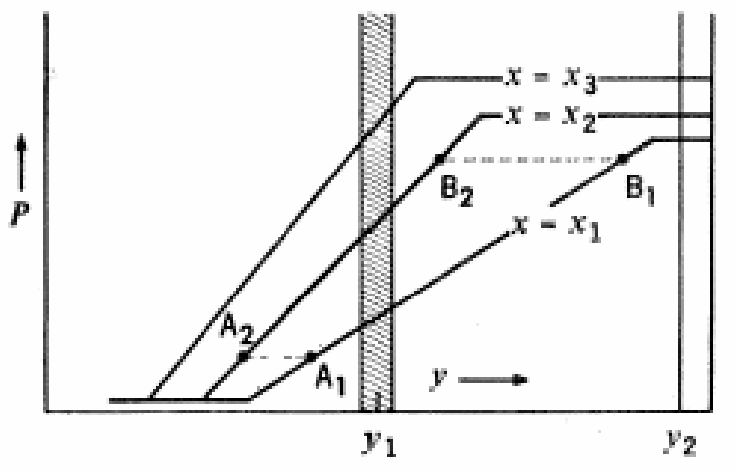

Figure 14. Layzer's hypothetical reaction norm graph for three genotypes (x_1, x_2, and x_3) exposed to a variable environment (y). (From Layzer 1974, Figure 1).

Layzer's project, like Lewontin's, was not entirely negative. Layzer also offered a positive thesis about what geneticists ought to be studying. And like Lewontin, Layzer pointed to the *causal mechanics* of the developmental genotype-environment-phenotype relationship. Layzer pointed to the developmental interventionist studies of Waddington. For example, in his first attack on Jensen, Layzer (1972a) drew on a quote from Waddington (1957) to contrast the "incomplete" calculation of heritability estimates with a more appropriate and "more penetrating" enterprise:

> …There has been a tendency to regard a refined statistical analysis of incomplete experiments as obviating the necessity to carry the experiments further and to design them in more penetrating fashion. For instance, if one takes some particular phenotypic character such as body weight or milk yield, one of the first steps in an analysis of its genetic basis should be to try to break down the underlying physiological systems into a number of more or less independent factors. Are some genes affecting the milk yield by increasing the quantity of secreting tissues, others by affecting the efficiency of secretion, and others in still other ways? (Waddington 1957, quoted in Layzer 1972a, 273)

Layzer was quick to contrast this enterprise with Jensen and Herrnstein's program, continuing, "These views contrast sharply with those of Jensen and Herrnstein, who believe in the possibility of discovering meaningful relations between measurable aspects of human behavior without inquiring into the biological or physiological significance of that behavior." (ibid) Layzer reiterated this positive thesis in 1974. After considering the limits of heritability estimates and their confinement to a specific mathematical theory, Layzer encouraged, "Other ways of assessing the effects of environment on phenotypically plastic traits may, however, be more useful in other contexts. In particular, certain kinds of intervention studies may provide more direct and more useful information about the effects of environment on IQ than conventional studies of IQ heritability." (Layzer 1974, 1260)

The conjoining of Lewontin and Layzer in Wagman's (1973) article, the citations of Layzer by Lewontin, and the citations of Lewontin by Layzer were no coincidence. Layzer, like Lewontin, emphasized the importance of using interventionist methodologies to study the causal mechanics of development and also the importance of considering possible, untested environments when discussing phenotypic variation and group differences. To unite these elements, Layzer, also like Lewontin, pointed to the concept of G×E, which revealed the complex interdependence of genetic and environmental factors and the fact that phenotypic variation due to these separate factors could change quite substantially with different genotypic or environmental distributions. Thus, Layzer, like Lewontin one more time, questioned the usefulness of the heritability estimates which overlooked this complex interdependence. Measuring the heritability of IQ, like measuring the sexuality of fractions, was closer to astrology and numerology than anything scientific.

3.2.2. The Argument from Genotype-Environment Interaction Dismissed

Lewontin and Feldman (1975) charged Jensen (1969) and Herrnstein (1971) with ignoring the cases of G×E on which they placed so much weight. In replying to Lewontin and Feldman, though, Jensen was quick to point out that he *had* considered G×E in his assessment (Jensen 1976). And indeed, Jensen had taken up the matter of G×E already in his 1969, although only to brush aside the complications that the concept might raise for his argument. There, Jensen bemoaned, "There is considerable confusion concerning the meaning of interaction in much of the literature on heredity and intelligence." (Jensen 1969, 39) He was critical of the growing group of "interactionists" that were merely masked environmentalists, writing, "Those who call themselves 'interactionists', with the conviction that they have thereby either solved or risen above the whole issue of the relative contributions of heredity and environment to individual differences in intelligence, are apparently unaware that the preponderance of evidence indicates that the interaction variance, V_I, is the smallest component of the total phenotypic variance." (ibid) Jensen, at this early time, was already anticipating criticisms of his genetic hypothesis with arguments from G×E: "The magnitude of V_I [i.e., variation due to G×E] for any given characteristic in any specified population is a matter for empirical study, not philosophic debate. If V_I turns out to constitute relatively small proportion of the total variance, as the evidence shows is the case for human intelligence, this is not a fault of the analysis of variance model. It is simply a fact. If the interaction variance actually exists in any significant amount, the model will reveal it." (ibid, 41) In his short discussion of G×E in 1969, Jensen introduced in an early form the basic arguments he would continue to employ when discussing G×E throughout the IQ controversy: (a) that invocations of 'interaction' were often simply confused about the meaning of the concept, (b) that discussions of G×E must be based on empirical data and not merely on

"philosophic" speculation, (c) that, if G×E existed for a particular phenotypic trait, then the analysis of variance would reveal such interaction, and finally (d) that the analyses of variance employed to measure sources of phenotypic variation for IQ revealed no such G×E.

Jensen employed the first of these arguments to criticize any invocation of developmental biology when discussing G×E. Under a section entitled "The Meaning and Non-meaning of 'Interaction'" in his *Educability and Group Differences* (1973), Jensen again grappled with his so-called 'interactionists': "Thus the interactionist theory holds that although there may be significant genetic differences at the time of conception, the organism's development involves such complex interactions with the environment that the genetic blueprint, so to speak, becomes completely hidden or obscured beneath an impenetrable overlay of environmental influences." (Jensen 1973, 49) Jensen explained that this interactionist position "has arisen from a failure to understand the real meaning of the term 'interaction' as it is used in population genetics; but even more it is the result of failure to distinguish between (a) the *development* of the individual organism, on the one hand, and (b) *differences* among individuals in the population." (ibid, emphases in original) Thus, any discussions of G×E that drew on the complexities of development were simply confused, and so there was no reason to engage with the confused arguments further.

Jensen employed the second argument against "philosophic" speculation when the question of the inherent locality of the analysis of variance and heritability estimates were criticized. "The methods of biometrical genetics, of course, have no power to predict h^2 under as yet untried interventions in the internal or external environments," he admitted when replying to "misconceptions about heritability." (Jensen 1975, 173) "It does give an indication of the relative influence of *existing* environmental sources of variance, and if h^2 is very high, it tells us that

merely reallocating individuals in existing environments will not have much effect in the rank ordering of individual differences." (ibid, emphasis in original) So the emphasis, for Jensen, was on the *actual* environments in existence and not on what the heritability of a trait might be in possible, as-yet-untested environments. Jensen drew on this point when replying to Lewontin's (1970a) first attack. Lewontin claimed that compensatory education could not be written off as a failure simply because it proved unsuccessful in the existing educational environments because new environments could be encountered or created that would facilitate such compensation. To this Jensen countered, "Lewontin seems to believe that anything is possible, given sufficient technological implementation. But reality does not bow to technology. Technology depends upon a correct assessment of reality." (Jensen 1970, 20) Lewontin's speculative emphases on *possible* environments were thus written off as beyond the domain of the empirical, quantitative studies of *actual* environments in which Jensen was interested. Herrnstein, likewise, considered and then dismissed such emphases on the implications of G×E in possible environments, writing, "It is, in fact, entirely possible that science could uncover ways of raising people's I.Q.'s by special sorts of environments, tailor-made for them. A world in which each person enjoyed something approaching his optimal environment—let us assume a different environment for each—might register large interaction and little overall variation in I.Q. That is, however, not our world, and we have as yet hardly any inkling of how to get from here to there, or even of whether or not the way exists in any practical sense." (Herrnstein 1973, 180)

When it came to the existence of G×E for IQ in actual environments, Jensen claimed that the analysis of variance was perfectly capable of detecting any such interactions, and so the fact that twin studies had not identified any such variation due to G×E revealed the lack of any such variation in nature. Responding to Layzer's (1972a) first commentary, Jensen noted, "Layzer

makes much of the possibility of interaction of genetic and environmental factors." (Jensen 1972, 435) But then he continued,

> The existing models of heritability analysis take such interaction into account and are capable of estimating the proportion of variance attributable to such interaction. With respect to IQ, the fact is that this interaction component is either nonexistent or so insignificant as to be undetectable in the existing data. If it were of substantial magnitude, it would easily show up with the present methods of analysis, which are quite capable of detecting other forms of interaction, such as dominance. In reading Layzer, one might easily get the impression that there is a lot of G×E interaction but that our models are unsuited to detecting it (ibid).

But Jensen responded, "Not so," citing Jinks and Fulker (1970), who conducted a survey of statistical approaches to the analysis of human behavior and concluded, "Unfortunately, an apparent lack of evidence of substantial genotype-environment interaction in intelligence-test scores strongly suggests that none of the range of environments provided by our society is likely uniformly to produce a high (or low) level of intelligence." (Jinks and Fulker 1970, 324) And again, Herrnstein echoed this conclusion, claiming that "the data from the twins reveal no interaction (in the technical sense) of heredity and environment." (Herrnstein 1973, 180)

With these arguments in mind, Jensen ultimately placed the burden of proof regarding the importance of G×E on those who wanted to emphasize the concept's importance: "If G×E interaction is held up as a criticism or limitation of the applicability of heritability analysis to mental test data, the burden of demonstrating the presence of substantial G×E interaction in such data must be assigned to the critics." (Jensen 1975, 182)

3.3. Science Studies on Genotype-Environment Interaction in the IQ Controversy

G×E, as evinced from the discussion in the previous section, figured prominently in the IQ controversy. Lewontin and Layzer took the matter of G×E to be not just important for

discussions of heritability but to fundamentally undermine the entire heritability enterprise. Ignorant of the causal mechanics of development or the phenotypic outcomes in possible environments, Lewontin and Layzer warned, the analysis of variance was useless. Instead, the focus of human geneticists should be on employing interventionist methodologies that reveal the causal mechanics of the developmental genotype-environment-phenotype relationship. Jensen and Herrnstein, however, were undeterred by the Lewontin-Layzer assault on heritability estimates. They countered by arguing that invocations of possible environments and developmental biology only confused discussions of individual differences and G×E. The analysis of variance, they responded, was perfectly capable of detecting the presence of any G×E; since it did not detect it in the actual populations examined, focus on such interaction was misleading.

As philosophers and historians of science have investigated the IQ controversy so too have they investigated G×E.[51] However, the tendency of such investigations has been to simply adopt one of the arguments on authority summarized in the paragraph above with an eye towards attacking the other, thus introducing the G×E debate into the domain of science studies. Sesardic (1993), remember, cited Jensen and Herrnstein to question the pervasiveness of G×E in nature. In his subsequent *Making Senses of Heritability* (2005), Sesardic also followed Jensen in distinguishing a 'technical' sense of G×E from a purportedly muddle-headed 'interactionism'. In Sesardic's terminology, there was a *statistical* notion of interaction (interaction_s) and a *commonsense* notion of interaction (interaction_c) (Sesardic 2005, 48). And, following Jensen again, Sesardic claimed that much confusion followed from invoking the individual-level,

[51] As mentioned in the previous section, Jensen's (1969) assessment was criticized from a variety of perspectives. In turn, historians, philosophers, and sociologists of science have converged on these various lines. Discussing each of these separate critiques would be beyond the scope of this essay, though, and so attention will be given only to those science studies scholars who discuss G×E. Again, Aby and McNamara (1990) provide a useful compilation of these other criticisms.

commonsense notion of interaction to criticize heritability estimates of individual differences, which were only complicated by the population-level, statistical notion of interaction: "Layzer's argument (defended by many authors) that complexities of developmental processes preclude the possibility of partitioning the phenotypic variation into genetic and environmental components seems to be the result of confusing different levels of analysis." (ibid, 73) And, following Jensen one more time, Sesardic explained that even statistical interaction rarely posed a problem for heritability estimates, since it could often be eliminated with a transformation of scale (ibid, 53).

Sesardic has not been alone in this attack on G×E and the accompanying norm of reaction. In his defense of *Why Race Matters* (1997), philosopher Michael Levin painted a similar picture and adopted similar arguments. Citing Lewontin and Layzer, Levin acknowledged that "The most popular reason for discounting genotypic differences, however, is genotype/environment interaction." (ibid, 229) Levin agreed that genotypes can express themselves differently in different environments, that two genotypes expressing themselves differently in one environment may express themselves identically in another environment, and that environmental manipulation might reduce a gap in IQ scores for black and white populations. However, Levin responded, "An obvious objection to this argument is that it very nearly treats what is possibly possible as if it were actual.…But the sheer *possibility* of environments in which the races agree in (nonzero) intelligence does not show that such environments actually exist, or that, if they do, they could sustain a human society. The reaction ranges for some genes, like those controlling eye color, is quite narrow; IQ may be equally constrained….Proponents of the interaction argument offer no evidence that their conjectures are more than that…(ibid, 229, emphasis in original)." So even though Lewontin (1970a) pointed to the social basis of current developmental environments generating a gap in IQ scores for black

and white populations and the need for considering future developmental environments that could eliminate this gap, Levin answered, "In short, even if all 'political' questions look forward to what can be done, gene/environment interaction shows at most that biology *might* be 'neutral,' not that it *is*. Should the reaction range for the IQ gene be narrow—and the mere possibility of its being wide proves nothing to the contrary—biology may forbid what 'social justice' demands." (ibid, 230, emphases in original)

Other historians and philosophers, however, have been much more charitable to arguments from G×E than Sesardic (1993, 2005) and Levin (1997). Indeed, the opposing trend has been to simply opt for the opposite position on G×E proposed in the IQ controversy, replacing Jensen and Herrnstein's *criticisms* of possible environments, developmental biology, and norms of reaction with Lewontin and Layzer's *emphases* on possible environments, developmental biology, and norms of reaction. Chase (1980), remember, took Lewontin and Layzer's discussion of interaction to undermine Jensen's very understanding of statistical interaction. In his *Intelligence and Race: The Origins and Dimensions of the IQ Controversy* (1979), Douglas Lee Eckberg also emphasized the importance of recognizing developmental biology and norms of reaction, writing, "For the individual genotype, development is specific to the environmental variables that it contacts, the time periods in which they are contacted, and the order in which they are contacted. The result is a unique norm of individual reaction that cannot be predicted in advance. The term *norm of reaction* indicates that phenotypic development is contingent on both the specific genotype and the specific environmental milieu contacted." (ibid, 90, emphasis in original) And Eckberg did not overlook the implications diagnosed by Lewontin and Layzer of being unable to generate norms of reaction for human populations:

> If genetic standardization is lacking, then the norms of reaction for the varied members of a given population will make it impossible to know what elements in the environment affect which members in

> what ways (Williams 1969). These problems appear not only in studies of wild-type populations, but also within the laboratory. For example, genotype-environment interaction is a problem that often affects agricultural applications of research (Feldman and Lewontin 1975, 1164)…With freely-mating populations, in which few controls on genotype exist at all—and for which only such crude devices as the statistical average of assortative mating are available—the usefulness of heritability is so eroded that a number of writers have insisted that such estimation is of no value whatsoever, or even that it cannot be properly accomplished at all… (ibid, 94).

Michael R. Matthews (1980), historian and philosopher of science education, echoed this emphasis on norms of reaction and also adopted the dichotomy between studying variation with the statistical analysis of variance and reaction norms in his philosophical examination of the implications of the IQ controversy for education. Criticizing the analysis of variance, Matthews drew on Lewontin (1974) to describe the implications of choosing a reaction norm approach:

> An important consequence of this change to norms-of-reaction research is that the analysis of variance, that time-honoured pursuit of IQ researchers, has next to no value. It is always measured for a specific environment and genotype distribution….In norms of reaction, we are concerned with the functional relation holding between genotype and environment and their expression in phenotype. This allows prediction of trait performance in situations of environmental change (governmental interventions). Richard Lewontin [(1974)] traverses this terrain of population genetics and its putative connections with IQ research, and concludes: 'The simple analysis of variance is useless for these purposes and indeed has no use at all. In view of the terrible mischief that has been done by confusing spatiotemporal local analysis of variance with the global analysis of causes, I suggest we stop the endless search for better methods of estimating useless qualities' (ibid, 146-147).

As the citations in this section reveal, the G×E debate has followed the lead of its scientific origins and entered the domain of science studies. But the science studies scholars have offered little in the way of any novel contribution to evaluating this debate over G×E. Instead, the trend has been simply to adopt the arguments of either Jensen and Herrnstein, or Lewontin

and Layzer in order to criticize the opposing position on the matter. With a strategy such as this, important historical and philosophical questions are inevitably left lingering: We were not told *why* the disparate assessments concerning G×E existed in the IQ controversy in the first place; the best that was offered in the way of an explanation for the disparity came from accusing one side of confusion or the other side of blinded dogma. Is this really the only explanation? Moreover, by simply adopting the arguments of the scientific authorities, we were not given any analysis of *how* the various components of these arguments congregated to generate the disparate conclusions found in the IQ controversy. And finally, we were not offered an evaluation of *where* the debate over G×E found in the IQ controversy fits into to the broader history of the concept; the debate over G×E was treated as an isolated event originating with Jensen, Herrnstein, Lewontin, and Layzer.

3.4. $G \times E_B$ vs. $G \times E_D$

The goal of this section is to begin answering these lingering questions with a new conceptual framework for discussing G×E in the IQ controversy. The thesis is that Jensen and Herrnstein, on the one hand, and Lewontin and Layzer, on the other hand, utilized two distinct concepts of G×E, each situated in its own respective research tradition: Jensen and Herrnstein utilized the *biometric* concept of G×E, or $G \times E_B$, while Lewontin and Layzer utilized the *developmental* concept of G×E, or $G \times E_D$.

Jensen and Herrnstein, like Fisher before them, were operating in the biometric tradition of biology. The problem on which they focused was the *relative contributions of nature and nurture to individual differences* in IQ. They sought to analyze the *causes of variation* responsible for these individual differences. They asked, *how much* of the variation in IQ is due

to differences in genotype and differences in environment? They attempted to partition the causes of variation responsible for these differences. And as Jensen said, "The mathematical technique for doing this, called 'the analysis of variance', was invented by Sir Ronald Fisher, the British geneticist and statistician. It is one of the great achievements in the development of statistical methodology." (Jensen 1969, 28)

However, as was seen in the last chapter, situating oneself in the biometric tradition and utilizing Fisher's analysis of variance had important implications for how one treated G×E. Answering the questions about how much variation for a phenotypic trait was due to *either* environment *or* heredity were easiest if the variation for the phenotypic trait was generated by *only* heritable *and* environmental variation. Variation due to G×E complicated this picture because it eliminated one from treating phenotypic variance simply as the sum of heritable and environmental variances, as Fisher readily admitted (Fisher and Mackenzie 1923). For example, if phenotypic variation was simply the sum of heritable and environmental sources of variation, then a study of identical twins reared apart in truly unique environments would measure environmental variance for a particular trait; this environmental variance could then be subtracted from the total phenotypic variance for the trait, and the remaining variance could be attributed to heritable variation. If there is variation due to G×E for the trait, however, then such a calculation would be inaccurate, for the calculated heritable variance would actually also include the variation due to G×E, thus incorrectly inflating this estimate. G×E thus created a potential complication for Jensen's biometric program, prompting Layzer to claim that, for Jensen, G×E arose "like an uninvited party guest." (Layzer 1972b, 471)

But, echoing Fisher, Jensen did not take this *potential* complication to be a *proven* problem. This was because Jensen understood G×E to be strictly a statistical measure and

detectable with Fisher's analysis of variance. Jensen, already in 1969, was quick to point out that the sources of phenotypic variation considered by the analysis of variance contained a component due to G×E, namely V_I (Jensen 1969, 34, see Eq. 3.1 above). To those critics who attacked the analysis of variance for assuming that all effects of heredity and environment were strictly additive, Jensen replied, "The presence of V_I in Equation [3.1] explicitly shows that the heredity × environment interaction is included in the analysis of variance model, and the contribution of V_I to the total variance may be estimated independently of the purely additive effects of heredity and environment." (ibid, 41) Jensen, operating in the biometric tradition, employed the *biometric* concept of G×E, or $G \times E_B$, defined as *a statistical measure of the breakdown in additivity between genotypic and environmental sources of variation, which is generated by a statistical methodology such as the analysis of variance* (Tabery Forthcoming).

Recognizing Jensen's employment of $G \times E_B$ helps to explain why Jensen so quickly dismissed Lewontin and Layzer's emphases on development when discussing G×E. Employing $G \times E_B$, Jensen treated G×E strictly as a statistical measure descriptive of individual differences in a population, so invoking G×E in a discussion of individual development, Jensen retaliated, arose from a "failure to understand the real meaning of the term 'interaction' as it is used in population genetics" (Jensen 1973, 49). As we saw above, Jensen reiterated this warning on a number of occasions, distinguishing actual interaction from the confused "interactionist theory." This distinction was by no means unique to Jensen in the 1970s. Behavioral geneticists Robert Plomin, John C. DeFries, and John C. Loehlin provided a similar warning in 1977: "Unfortunately, discussions of genotype-environment interaction have often confused the population concept with that of individual development. It is important at the outset to distinguish genotype-environment interaction from what we shall call *interactionism*, the view

that environmental and genetic threads in the fabric of behavior are so tightly interwoven that they are indistinguishable." (Plomin, DeFries, and Loehlin 1977, 309) This distinction between statistical interaction and a muddle-headed interactionism has become standard fare for quantitative behavioral geneticists (Bouchard and Segal 1985, 393; Plomin and Hershberger 1991, 31; Scarr 1995, 155-157; Surbey 1994, 263-264). And, as was shown above, Sesardic (2005) also appropriated it with his distinction between interaction_s and interaction_c.

Recognizing Jensen's employment of $G\times E_B$ also explains why Jensen was undeterred by Lewontin and Layzer's emphases on possible environments when discussing G×E. When pressed on the locality of the heritability estimates, Jensen simply confined his genetic hypothesis to the actual environments. As far as Jensen was concerned, Lewontin and Layzer could speculate all they wanted about what *might* happen in as-yet-untested environments, but this speculation would not change what was *actually* occurring in the tested environments. And again, Jensen was not alone in making this point during the IQ controversy. Plomin and DeFries, this time directly responding to Feldman and Lewontin (1975), wrote,

> …Feldman and Lewontin reiterate the common knowledge that heritability estimates are limited to the population sampled and that genotype-environment interaction and correlation may be important. These points are misinterpreted by Feldman and Lewontin to mean that quantitative genetic analyses are, therefore, of no use. The conclusion does not follow (Plomin, DeFries, and Loehlin 1977). The very purpose of quantitative genetic studies is to *describe* genetic variability *in a specific population* and to ascribe that variability to environmental differences and genetic differences in that population (Lush 1940)….Feldman and Lewontin seem to be more concerned with the question of *what could be* rather than *what is*. (Plomin and DeFries 1976, 11, emphases in original).

In summary, with $G\times E_B$ situated in the traditional biometric tradition focused on the problem of estimating the relative contributions of heredity and environment, and with $G\times E_B$ understood to be a strictly statistical measure detectable with the statistical analysis of variance

(see Table 5. below), Jensen and Herrnstein simply dismissed Lewontin and Layzer's invocations of development and possible environments in discussions of $G \times E_B$ as confused and misleading. Ultimately, Jensen and Herrnstein fell into line with Fisher by claiming that the burden of proof rested with the critics of the genetic hypothesis to show that interaction was anything more than a *possible* complication for the heritability estimates of IQ.

Components	**Biometric Research Tradition**
Problem	Individual Differences
Approach to Causation	Causes of Variation
Causal Question	How Much?
Methodology	Statistical
Concept of Interaction	Biometric—$G \times E_B$

Table 5. The components of the biometric research tradition.

In contrast, Lewontin and Layzer, like Hogben before them, were operating in the developmental tradition of biology. They focused on the problem of unraveling the way in which variation in a population arose from the relationship between genotype and environment during *individual development*. They sought to analyze the *causal mechanics* of this relationship. They asked, *how* do differences in genotype and differences in environment relate during individual development to generate differences in phenotype? And they endorsed *interventionist methodologies*, such as those undertaken by Dobzhansky and Waddington.

Remember that Lewontin concluded his letter of 2 May 1973 to Dobzhansky with the complaint, "Why is it that most geneticists do not understand that the phenotype is a developmental process?" Understanding the phenotype to be a developmental process obviously had important implications for how Lewontin conceptualized the genotype-environment-

phenotype relationship. The genotype, the environment, and the phenotype could not be treated as individual units, according to Lewontin, the first adding to the second to create the third. Rather, the first and the second interacted continuously throughout development, and the third was the manifestation of this interactive, developmental process. Differences in the phenotype, then, would result from differences in this interactive, developmental process. Layzer's discussion of the phenotypic information processing skills loosely measured by IQ revealed a similar understanding: "Information-processing skills, like other skills," Lazyer explained, "are not innate, but develop over the course of time." (Layzer 1972a, 280) Like the complexities involved in the building of a house, Layzer argued that cognitive development was simultaneously influenced by genetic factors and the environmental challenges in which that development took place. Lewontin and Layzer pointed to G×E so often because it was seen to reflect this developmental genotype-environment-phenotype relationship. Extreme cases of G×E showed that even slight differences in genotypic or environmental distribution could lead to huge differences in phenotypic outcome, thus revealing the interdependence of the factors in this relationship.

Not surprisingly, then, Lewontin and Layzer suggested investigating G×E with interventionist methodologies that could appropriately dissect and display this developmental genotype-environment-phenotype relationship. Layzer pointed to the studies of Waddington that attempted to manipulate the developmental process in order to understand how these various components were interrelated. Lewontin referred to the reaction norm studies of Dobzhansky that placed different genotypic groups in different environments and then tracked the differences in viability of the different *Drosophila* strains in different environmental conditions (temperature).

Lewontin and Layzer, operating in the developmental tradition employed the *developmental* concept of G×E, or G×E_D, defined as *variation that results from differences in unique, developmental combinations of genotype and environment* (Tabery, Forthcoming). Recognizing Lewontin and Layzer's employment of G×E_D clearly reveals why they placed so much emphasis on the importance of considering *development* when discussing phenotypic variation and group differences. Furthermore, recognizing Lewontin and Layzer's employment of G×E_D also helps explain why they both placed so much emphasis on the importance of considering *possible* environments when discussing these issues. They advocated employing interventionist methodologies to investigate the causal mechanisms of individual development. From this point of view, possible environments suggested different values that could be taken by the variables in the mechanism. Unusual environments, even those that an organism may never encounter, represented a powerful interventionist tool on which to alter the values of the environmental variable and test the proposed mechanism (Griffiths and Tabery, Forthcoming).

In summary, with G×E_D taken to reflect the developmental genotype-environment-phenotype relationship, and with G×E_D understood to be properly investigated with methodologies that could dissect and display this relationship, Lewontin and Layzer took the consideration of both development and possible environments to be essential to any discussion of the heritability of IQ. Because the analysis of variance and the derivative heritability estimates omitted these essential elements of G×E_D, the statistical methodologies were useless. Following Hogben, Lewontin and Layzer placed the burden of proof on those geneticists and psychologists who assumed the additivity of genotype and environment was an accurate reflection of the biological reality (see Table 6).

Components	**Developmental Research Tradition**
Problem	Individual Development
Approach to Causation	Causal Mechanisms
Causal Question	How?
Methodology	Interventionist
Concept of Interaction	Developmental—$G \times E_D$

Table 6. The components of the developmental research tradition.

Recognizing the distinction between $G \times E_B$ and $G \times E_D$ along with the other components of each research tradition provides answers to the historical and philosophical questions that I have suggested linger if we evaluate the debate over G×E in the IQ controversy simply by adopting the arguments put forth by Jensen, Herrnstein, Lewontin, and Layzer. We can now understand why the disparate assessments of G×E's importance existed in the IQ controversy in the first place. Employing $G \times E_D$, Lewontin and Layzer placed much weight on the concept, for it suggested that untested, developmental environments might be encountered or created that could significantly eliminate the "achievement gap" between black and white populations. Employing $G \times E_B$, however, Jensen and Herrnstein criticized such emphases on development and possible environments as confused and misleading; analyses of variance would find any such interaction, they countered, and as long as they did not, appeals to this interaction as a means to criticize heritability estimates were only speculative.

Were Jensen and Herrnstein simply confused about the concept of interaction, as Chase (1980) asserted? Or were Lewontin and Layzer simply blinded by dogma concerning the importance of G×E and its depiction with reaction norms, as Sesardic (1993) asserted? Simply writing off Jensen and Herrnstein as confused or Lewontin and Layzer as blinded by dogma overlooks the fact that both sides of this debate, as section 3.2 displayed, offered multiple

reasons for their conclusions about the importance of G×E and also referenced the existing, limited, empirical data in support of their conclusion. Understood as a developmental concept, Lewontin and Layzer emphasized the importance of $G \times E_D$ because it reflected the nature of the developmental genotype-environment-phenotype relationship; moreover, it suggested that even slight alterations to this relationship (such as the slight modification of the environment) could have a huge impact on total phenotypic variation. So Lewontin and Layzer argued that a statistical methodology such as the analysis of variance and its derivative heritability estimates, which ignored the causal mechanics of this developmental relationship, was inevitably prone to misleading conclusions about how the genotype, the environment, and the phenotype were interrelated. But understanding G×E simply as a statistical measure generated by the analysis of variance, Jensen and Herrnstein dismissed this reference to development. Instead, they pointed to the statistical methodologies that measured variance due to interaction (i.e., V_I) to justify the minimal influence of this source of variation on the total phenotypic variation for IQ in a population. This epistemological divide can be seen most clearly when we view their separate research traditions side-by-side (see Table 7).

Components	**Biometric Tradition**	**Developmental Tradition**
Problem	Individual Differences	Individual Development
Approach to Causation	Causes of Variation	Causal Mechanisms
Causal Question	How Much?	How?
Methodology	Statistical	Interventionist
Concept of Interaction	Biometric—$G \times E_B$	Developmental—$G \times E_D$

Table 7. The components of the biometric and developmental research traditions.

The real dilemma was not that Jensen and Herrnstein were *confused*, nor that Lewontin and Layzer were *blinded*. The problem was the same that plagued Fisher and Hogben—there was just so *little* empirical data at the time that was utilizable in defense of either set of conclusions concerning the importance of G×E. Lewontin referenced the reaction norm research of Dobzhansky on *Drosophila* viability, but he also readily admitted that finding G×E with these reaction norms in humans would be impossible, and so *hypothetical* reaction norms were generally Lewontin's ammunition when discussing variation in human populations. Jensen and Herrnstein pointed to the twin studies existent at the time, but they also readily admitted that these were limited by low statistical power and the fact that separated twins were still often reared in relatively similar environments. And so the debate was inevitably one about what one should *assume* to be occurring in nature when faced with so little empirical data. Employing $G\times E_D$, Lewontin and Layzer pointed to the nature of development to argue for assuming that the interaction was standard. But employing $G\times E_B$, Jensen and Layzer pointed to the apparent ability of the additive model to fit the statistical data to argue for assuming that the interaction was minimal.

3.5. Conclusion

This chapter examined the place of G×E in the IQ controversy. Philosophers and historians of science who have focused on this topic have tended to do so by simply adopting on authority either Jensen and Herrnstein's arguments *against* or Lewontin and Layzer's arguments *for* the importance of G×E with regards to the heritability of IQ. I claimed that such an approach inevitably left important questions lingering concerning the existence and nature of these disparate assessments of G×E, along with the place of this debate within the broader history of

the concept. In contrast to this approach, I introduced a new conceptual framework for discussing G×E in the IQ controversy by distinguishing the *biometric* concept of G×E, or $G \times E_B$, and the *developmental* concept of G×E, or $G \times E_D$. Recognizing Jensen and Herrnstein's utilization of $G \times E_B$ and Lewontin and Layzer's utilization of $G \times E_D$ allows one to realize why the two sides came to such disparate conclusions concerning the importance of G×E and also realize why the two sides were able to give multiple, interrelated reasons for coming to these disparate conclusions.

4. DIFFERENCE MECHANISMS

Abstract. In recent years, philosophers of science have found a renewed interest in mechanisms. The motivation is the thought that the elucidation of a mechanism generates a causal explanation for the phenomenon under investigation. For example, a question such as, How do rats form spatial memories of their environments?, is answered by elucidating the *regular causal mechanisms* responsible for the *individual development* of spatial memory in rats. But consider a slightly different question: How do some rats come to have better spatial memory than other rats? This is a question about the *causes of variation* responsible for *individual differences* in spatial memory. The first question demands an answer about *regularity*; the second question demands an answer about *variation*. The account of causal-mechanical explanation on offer by philosophers of science captures regularity, but it neglects variation. In this chapter I attempt to modify the mechanical program so as to incorporate both regularity and variation. The task is to explicate the relationship between the regular causal mechanisms responsible for individual development and the causes of variation responsible for individual differences; the common denominator between the two is what I will call *difference mechanisms*. As it turns out, this is precisely the relationship that has divided the biometric research tradition and the developmental research tradition in the long-standing debates over genotype-environment interaction, or G×E. Ultimately, then, the product will be a modified account of causal-mechanical explanation that captures both regularity and variation, and which may be utilized to resolve the debates over G×E.

4.1. Introduction

In recent years, philosophers of science have found a renewed interest in mechanisms.[52] Unsatisfied with traditional law-based accounts of explanation which do not capture the nature of explanation in special sciences such as biology, philosophers have turned to mechanisms as an alternative. The motivation is the thought that the elucidation of a mechanism generates a causal explanation for the phenomenon under investigation (Bechtel and Richardson 1993; Glennan 2002; Machamer, Darden, and Craver 2000; Schaffner 1993; Woodward 2002). There are differences between the various accounts of a mechanism.[53] But the accounts hold in common

[52] I say "renewed" to contrast these more recent endeavors with the classical mechanical philosophy. On the virtues and vices of the 17th C. mechanical philosophy, see Dijksterhuis (1961, section 4), Gillispie (1960, chapter 3), Hall (1952), Westfall (1971), Wilson (1999).

[53] For instance, there is a difference in the way in which the parts of a mechanism are understood to behave. This behavior has been characterized as a *function* (Bechtel and Abrahamsen 2005), an *activity* (Machamer, Darden, and

the basic idea that a scientist provides a successful causal explanation by identifying and manipulating variables in a regular causal mechanism thereby determining how those variables are situated in and make a difference in the mechanism; the ultimate explanation then amounts to the elucidation of how those variables act and interact to produce the phenomenon under investigation. The accounts are meant to explain how scientists answer questions such as the following: How are neural messages relayed across a synapse (Machamer, Darden, and Craver 2000)? How do immune systems identify and attack antigens (Schaffner 1993)? How do plants convert solar energy into chemical energy (Tabery 2004)? How does *E. coli* determine whether or not to produce lactose-metabolizing enzymes (Woodward 2002)? Or, how do rats come to form spatial memories of their environments (Craver and Darden 2001)? The thought is that such questions are answered by elucidating the regular causal mechanisms responsible for synapse transmission, immunologic response, photosynthesis, gene expression, or the formation of spatial memory. To take just one example (see Figure 15), Carl Craver and Lindley Darden show how scientists provide an explanation for the phenomenon of spatial memory by elucidating the various entities and activities involved in regular causal mechanisms at the molecular, cellular, brain-system, and organismal levels to produce spatial memory in rats (Craver and Darden 2001).

Craver 2000), an *interaction* (Glennan 2002; Woodward 2002), and an *interactivity* (Tabery 2004). See Tabery (2004) for an analysis of this difference and the relationship between the various accounts.

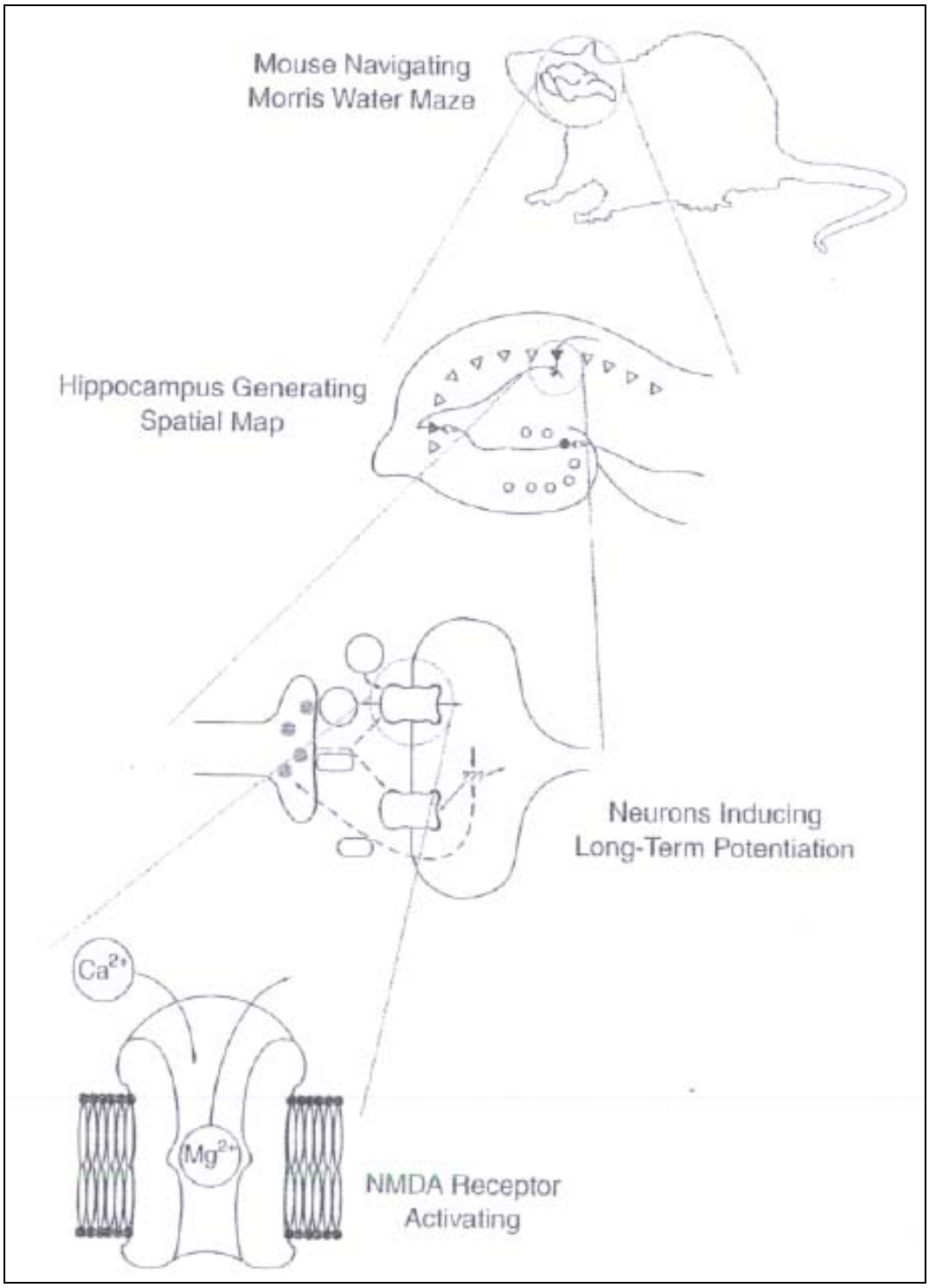

Figure 15. Molecular, cellular, brain-system, and organismal mechanisms involved in the production of spatial memory in rats. From Craver and Darden (2001, 6.4).

But consider a slightly different question: How do some rats come to have better spatial memory than other rats? One common way to test spatial memory in rats is to place them in a Hebb-Williams maze test; rats start at one corner and attempt to navigate to a food source at the other end (Hebb and Williams 1946). The rats, after attempting several practice mazes, then attempt test mazes; they are scored for how many "errors" they make, which occur when the rats deviate from the correct path by crossing one of the dotted lines (Figure 16). Not surprisingly, there is quite a bit of variation in how well different rats do in these maze tests; that is, some rats develop better spatial memory than other rats (Cooper and Zubek 1958). So: How do some rats come to develop better spatial memory than other rats?

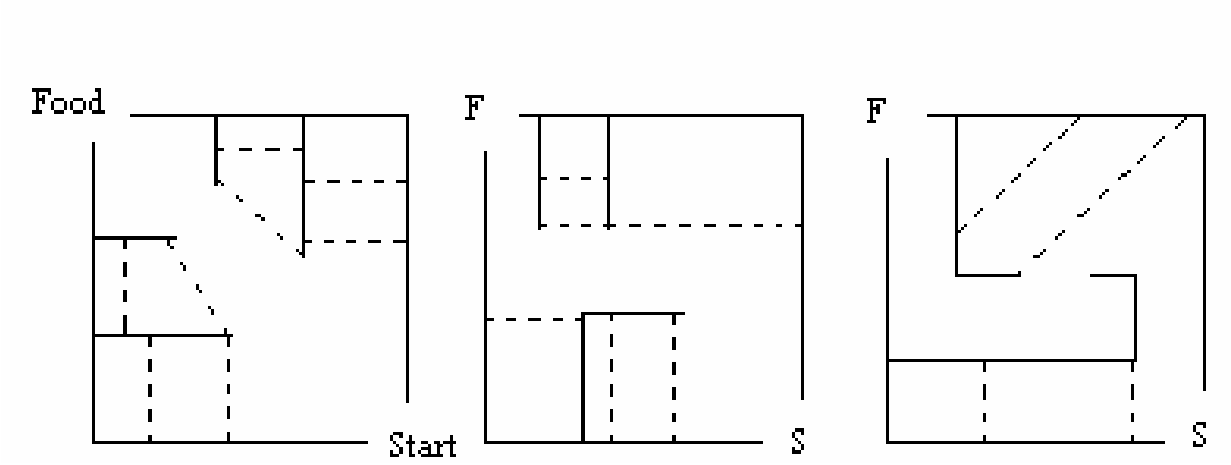

Figure16. Sample Hebb-Williams maze test configurations.

The question that Craver and Darden (2001) consider involves an investigation into the *regular causal mechanisms* responsible for the *individual development* of spatial memory. But the second question involves an investigation into the *causes of variation* responsible for *individual differences* in spatial memory. The first question demands and answer about

regularity. The second question demands an answer about *variation*. In focusing only on the first question, philosophers of science have thus far confined themselves to providing an account of causal-mechanical explanation that captures regularity. But what about an account that captures variation?

When you think about it, it is actually quite striking that the philosophers of science who turned to mechanisms left out variation. Variety is the spice of life, and variation is the space of natural selection. In contrast to a physicist who is generally safe assuming that an electron, is an electron, is an electron, a biologist is often interested in precisely what makes one species *different* from another, one population *different* from another, or one individual *different* from another because it is the *difference* that provides for the variation.[54] Philosophers of science, remember, have turned to mechanisms out of dissatisfaction with the accounts of scientific explanation that were forged in physics and did not apply to biology. And yet, one of the core features of the biological world—variation—is lacking from the mechanical program.

The purpose of this chapter is to modify the mechanical program in order to capture this essential element of biological explanation. The task, then, is to determine what the relationship is between, on the one hand, *regular causal mechanisms responsible for individual development* and, on the other hand, *causes of variation responsible for individual differences*.

As it turns out, formulating this relationship is not simply an exercise for the armchair philosopher. For it is precisely this relationship that resides at the heart of the nature-nurture debate. More specifically, it is precisely this relationship that resides at the heart of debates over genotype-environment interaction, or G×E, which were traced in the previous chapters. Biologists in both the biometric and the developmental traditions converge on the question of

[54] Obviously physicists are not *always* safe making this assumption; they too must heed potential variation (Hamilton 2006). But certainly not to the extent that biologists do.

G×E. Biologists in the biometric tradition employ statistical methodologies in order to answer *how-much?* questions about the causes of variation responsible for individual differences. Biologists in the developmental tradition employ interventionist methodologies in order to answer *how?* questions about the causal mechanisms responsible for individual development (Table 8). In the face of a debate that has raged now for nearly a century, a form of *isolationist pluralism* has emerged, wherein disputants from both sides have attempted to calm the storm by suggesting that these two traditions simply operate at different *levels of analysis*. The biometrically-oriented biologists investigate the causes of variation responsible for individual differences, and the developmentally-oriented biologists investigate the causal mechanisms responsible for individual development…and ne'er the twain shall meet. In this chapter, I want to challenge this isolationist pluralism with an *integrative* alternative. In so doing, I will use this as a case to develop the relationship between the regular causal mechanisms responsible for individual development and the causes of variation responsible for individual differences, thereby generating the modified account of causal-mechanical explanation that captures both regularity and variation.

Components	**Biometric Tradition**	**Developmental Tradition**
Problem	Individual Differences	Individual Development
Approach to Causation	Causes of Variation	Causal Mechanisms
Causal Question	How Much?	How?
Methodology	Statistical	Interventionist

Table 8. The components of the biometric and developmental research traditions.

The thesis will be the following: The biometric research tradition and the developmental research tradition may be united based upon a shared problem—the elucidation of what I will call *difference mechanisms*. Differences mechanisms are regular causal mechanisms made up of difference-making variables that take different values in the natural world. I will have to unpack this basic idea by explicating the more specific relationships between the various components of each tradition—(a) individual differences vs. individual development, (b) causes of variation vs. causal mechanisms, (c) how-much? vs. how? causal questions, and (d) statistical vs. interventionist methodologies. I will attempt this unpacking by drawing on and extending the idea of understanding causes as *difference-makers*, which has been developed by several philosophers in recent years (Lewis 1973; Woodward 2003). The relationships, then, will look like this: (a) Individual differences are the effect of *difference-makers in development that take different values in the natural world*, or, in C. Kenneth Waters' (Forthcoming) locution, when the variables are *actual* difference-makers; (b) the difference-making variables in the regular causal mechanisms responsible for individual development simultaneously are the causes of variation when the difference-making variables naturally take different values; (c) how-much? and how? are various causal questions that one may ask about these difference-makers in development; and (d) statistical and interventionist methodologies are both tools that may be used to investigate theses difference-makers in development. Finally, I will take this general framework and apply it to the debate over G×E, arguing that G×E results from the *interdependence of difference-makers in development that take different values in the natural world*. Ultimately, then, the product will be a modified account of causal-mechanical explanation that captures both regularity *and* variation, and which may be utilized to resolve the debates over G×E.

4.2. The Case of Genotype-Environment Interaction

It is a truism that genes and the environment interact during the course of individual development. Scientists in the biometric tradition such as quantitative behavioral geneticists, however, traditionally ask questions about *how much causes of variation are responsible for individual differences*, not questions about *how regular causal mechanisms are responsible for individual development*. For example, authors of the popular *Behavioral Genetics* textbook write, "For the complex traits that interest psychologists, it is possible to ask not only *whether* genetic influence is important but also *how much* genetics contributes to the trait. … The question about how much genetics contributes to a trait refers to *effect size*, the extent to which individual differences for the trait in the population can be accounted for by genetic differences among individuals." (Plomin et al. 1997, 77-78) For scientists in the biometric tradition, the problem of interest is not the causal-mechanical interplay between genes and the environment during, for example, gene expression or synapse formation; the problem of interest is the relative contributions of genetic differences and environmental differences to individual differences for a trait in a population.

The standard methodology for investigating individual differences is the statistical analysis of variance (ANOVA). In its simplest form, ANOVA partitions total phenotypic variation (V_P) into a source attributable to genetic variation (V_G) and a source attributable to environmental variation (V_E):

$$V_P = V_G + V_E \quad (4.1)$$

In this simplest of cases, the two sources of variation are additive, meaning that V_G and V_E (the "main effects") together fully account for V_P. When this simple case applies, we can also then

talk about the proportion of the total variation attributable to genetic or environmental differences; for example, the proportion of genetic variation is referred to as the *broad heritability* (h^2) of a trait, calculated as:

$$h^2 = V_G/V_P \quad (4.2)$$

However, when different genetic groups respond differently to the same array of environments, the additivity between V_G and V_E breaks down, requiring an addition to the equation in the form of G×E. G×E creates a potential problem for biometricians because it generates its own variation ($V_{G\times E}$), breaking down the additivity in Equation (4.1), forcing a modification to Equation (4.3), and also eliminating the ability to calculate the heritability of a trait unless a transformation of scale is employed to make the variation due to G×E disappear, wherein the scale on which the variables are measured is changed in order to get back to an additive relationship between the main effects.

$$V_P = V_G + V_E + V_{G\times E} \quad (4.3)$$

The additive and non-additive situations in Equations (4.1) and (4.3) may also be contrasted by considering reaction norm graphs, such as those in Figures 17(A) and 17(B). Three genetic groups are represented in the graphs, each with its own reaction norm.[55] The three groups are differentiated based on the particular variant of the promoter region in the serotonin transporter gene (5-HTT) carried (s/s vs. s/l vs. l/l) and measured for a particular trait (probability of a major depression episode, y-axis) across an array of environments (number of stressful life events experienced, x-axis). The "s" stands for a *short* form of the promoter region in the gene, while the "l" stands for a *long* form of the promoter region. Individuals receive either an *s* or an *l* from each of their parents, and the *short* promoter region generates relatively

[55] For a history of the reaction norm concept, see Sarkar (1999). For a comparison of the reaction *norm* concept with the reaction *range* concept, see Griffiths and Tabery (Forthcoming).

less serotonin binding than the *long* promoter region. When V_G and V_E are additive, then the reaction norms will be parallel as they are in the hypothetical example found in Figure 17(A). But when V_G and V_E are not additive—when there is G×E, then the reaction norms will be non-parallel as they are in Figure 17(B) drawn from empirical data (Caspi et al. 2003).

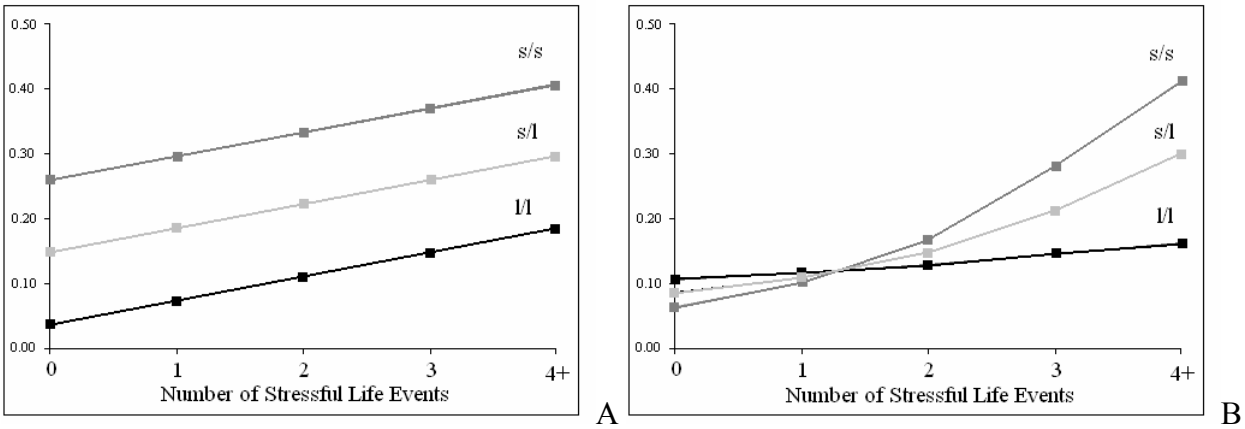

Figure 17. Reaction norm graphs for s/s, s/l, and l/l groups measured for probability of a major depression episode (y-axis) across an array of environments (x-axis). (A) Hypothetically parallel reaction norms. (B) Non-parallel reaction norms drawn from empirical data (Caspi et al. 2003).

4.2.1. The Call for an Isolationist Pluralism

Notice that we have now introduced two notions of interaction: (a) the interaction between genes and the environment in the regular causal mechanisms responsible for individual development, and (b) the interaction between genetic and environmental causes of variation responsible for individual differences in a population. What does the former notion of interaction have to do with the latter? Scientists in the developmental tradition criticize scientists in the biometric tradition for their focus on main effects and their attempts to avoid the complications posed by

G×E. More specifically, they understand G×E to somehow reflect the developmental relationship between genes and the environment, and so ANOVA's trouble with G×E, they argue, is symptomatic of the biometricians' more general trouble with elucidating the causal mechanisms of individual development. For example, David Layzer warns, "For complex animal characters there is little reason to expect additivity and independence to prevail. On the contrary, such characters usually reflect a complicated developmental process in which genetic and environmental factors are inextricably mingled." (Layzer 1972, 275; see also Layzer 1974) And Richard Lewontin scolds, "…relations between genotype, environment, and phenotype are at base mechanical questions of enzyme activity, protein synthesis, developmental movements, and paths of nerve conduction. We wish, both for the sake of understanding and prediction, to draw up the blueprints of this machinery and make tables of its operating characteristics with different inputs and in different milieus. For these problems, statistical descriptions, especially one-dimensional descriptions like heritability, can only be poor, and, worse, misleading substitutes for pictures of the machinery." (Feldman and Lewontin 1975, 1167-1168; see also Lewontin 1974) Michael Meaney writes, "The cellular context, and specifically levels of transcription factor such as cFos and cJun, are heavily influenced by ongoing activity; stress, social encounters—all serve to influence the cellular levels of these factors and can therefore have very potent influences on the nature of gene activity. From such systems will we derive main effects? I think not." (Meaney 2001, 53) Gilbert Gottlieb also complains, "The population view of behavioral genetics is not developmental. It is based on the erroneous assumption that a quantitative analysis of the genetic and environmental contributions to individual differences sheds light on the developmental process of individuals." (Gottlieb 2003, 338) And G. J. Vreeke concurs, "An analysis of variance abstracts from (actual) interaction effects and thus cannot offer

an accurate picture of development…. Behavioral geneticists, then, should acknowledge that an analysis of variance is a statistical method that does not fit reality and should be judged against the background of the best material model we have of development, which is one of dynamics and interactions." (Vreeke 2005, 44)

This critical analysis of the biometric tradition, and especially that formulated by Lewontin and Layzer, resonates in the philosophy of science. Philosophers including Block and Dworkin (1976), Block (1995), Daniels (1974), Downes (2004), Kaplan (2000), Sarkar (1998), and Sober (1984) all reiterate Lewontin and Layzer's criticisms of ANOVA, emphasizing the statistical methodology's trouble with G×E along with its inability to elucidate the causal mechanisms of individual development.

Faced with such criticisms based on the importance of interaction, scientists in the biometric tradition naturally mount a defense. To do so, they draw a firm distinction between their focus on individual differences along with the statistical concept of G×E that applies to individual differences, and the causal-mechanically-minded biologists' focus on individual development along with their developmental notion of interaction. Robert Plomin, John DeFries, and John Loehlin, for instance, begin their assessment of G×E by complaining, "Unfortunately, discussions of genotype-environment interaction have often confused the population concept with that of individual development. It is important at the outset to distinguish genotype-environment interaction from what we shall call *interactionism*, the view that environmental and genetic threads in the fabric of behavior are so tightly interwoven that they are indistinguishable (Plomin, DeFries, and Loehlin 1977, 309). This distinction between the population-level concept of G×E and a purportedly muddle-headed "interactionism" is often deployed by defenders of ANOVA. Arthur Jensen, as we saw in Chapter 3, utilizes the defense in an attempt to protect his

employment of ANOVA and heritability estimates for attributing the gap in average IQ scores between black and white populations to genetic differences (Jensen 1969). So when critics such as Layzer and Lewontin appeal to the interaction between genes and the environment during the course of individual development, Jensen replies that "'interactionism' has become merely a substitute for extreme environmentalism. … This position has arisen from a failure to understand the real meaning of the term 'interaction' as it is used in population genetics; but even more it is the result of failure to distinguish between (a) the *development* of the individual organism, on the one hand, and (b) *differences* among individuals in the population." (Jensen 1973, 49)

This distinction between statistical interaction and a muddle-headed interactionism is often framed in terms of a difference in *levels of analysis*.[56] For instance, Thomas Bouchard and Nancy Segal complain, "It is common for theorists of the heredity × environment controversy to confuse the statistical concept of interaction with a viewpoint called interactionism. The problem arises because each concept applies at a different level of analysis." (Bouchard and Segal 1985, 393) And Michele Surbey, responding to a critique of the biometric tradition based on G×E, responds, "…the level of analysis at which [quantitative behavioral geneticists] are working is relatively insensitive to interactions. The concept of heritability describes characteristics of a population while the examination of ontogenetic interactions occurs at a distinctly different level of analysis." (Surbey 1994, 263)

The distinction between interaction and interactionism conjoined with this notion of different levels of analysis now also penetrates the philosophy of science. Neven Sesardic (2005)

[56] As Craver (2001, 2002, 2005), Schaffner (1993), and Wimsatt (1972, 1976, 1984) helpfully show, the concept of "levels" is utilized by scientists and philosophers in a variety of fashions—levels of aggregation, levels of organization, levels of analysis, levels of explanation. In the nature-nurture debate, the common choice is "levels of analysis," although the distinction that is being referenced is that between the population level and the individual level, which would be more accurately referred to as a part-whole relationship capturing different levels of organization.

also distinguishes two forms of interaction: commonsense interaction ($interaction_c$) and statistical interaction ($interaction_s$). "$Interaction_c$ of genes and environments is always present but it generates no problems for the estimation of heritability," Sesardic claims. "On the other hand, the existence of strong $interaction_s$ between genes and environments may really undermine the usefulness of heritability claims, yet the existence of such interaction is itself an open empirical question. Briefly, $interaction_c$ is ubiquitous but irrelevant for discussions about heritability, whereas strong $interaction_s$ is potentially a problem for heritability, but the extent of its presence remains a contentious issue." (Sesardic 2005, 49) So in response to Layzer, who criticizes Jensen for ignoring the complications posed by individual development, Sesardic counters, "Layzer's argument (defended by many other authors) that complexities of developmental processes preclude the possibility of partitioning the phenotypic variation into genetic and environmental components seems to be the result of confusing different levels of analysis." (Sesardic 2005, 73)

An appeal to different levels of analysis is by no means unique to the nature-nurture debate. There is a long history in biology of uniting different explanation-seeking *questions* with different explanatory *levels* (Huxley 1916; Mayr 1961[57], 1982; Tinbergen 1963). And appealing to these different levels or questions as a means of defense against cross-disciplinary criticism is also by no means unique to the nature-nurture debate. Paul Sherman (1988, 1989) argues that debates over the evolutionary origins of the female orgasm result from confusing different levels of analysis. Developmental biologists seeking answers to questions about how the clitoris develops, Sherman claims, can not use their results to attack evolutionary biologists seeking answers to questions about reproductive success. Interestingly, Sherman's appeal to different levels of analysis also bears other striking similarities to the appeal in the nature-nurture debate. Just as Plomin, Jensen, and Sesardic claim that their critics suffer from confusion over the

[57] For analyses of Mayr's famous paper, see Ariew (2003) and Beatty (1994).

multiple meanings of "interaction," so too does Sherman attempt to explain the female orgasm debate by claiming that there is confusion over the multiple meanings of "adaptation." And just as the defenders of ANOVA claim that their critics fail to distinguish between biometric *how-much?* questions and developmental *how?* questions, so too does Sherman attempt to explain the female orgasm debate by claiming that there is a failure to distinguish between developmental *how?* questions and evolutionary *why?* questions.

The appeal to different levels of analysis in the nature-nurture debate and the female orgasm debate both affirm a type of *isolationist pluralism* (Mitchell 2003). Biologists in the biometric tradition and biologists in the developmental tradition (or developmental biologists and evolutionary biologists in the female orgasm debate) focus on different problems; they employ different causal approaches; they ask different causal questions; and they utilize different methodologies. Indeed, the whole point of Chapters 2 and 3 was to tease apart these different components of the biometric and the developmental traditions, as was shown earlier in Table 8. Does it follow, however, that the two traditions cannot inform one another because of these differences? Does it follow that they pluralistically co-exist, but only by virtue of explanatory closure at each level? The appeal to different levels of analysis answers "Yes" to both of these questions.

But isn't this an odd position in which to find oneself—to think that scientists from different traditions can converge on a common phenomenon, such as complex behavioral traits or the female orgasm, and yet believe that these scientists have nothing to say to each other simply because they are asking different questions about the phenomenon? Sandra Mitchell (1992, 2002, 2003) asks just this question of Sherman's analysis of the female orgasm debate. And different questions, she explains, need not necessarily isolate scientists from critically

engaging each other. In contrast, she promotes an *integrative pluralism*: "Thus, although pluralism within biology is to be applauded, it is not the pluralism of questions and the consequent independence of answers that Sherman endorses, but rather a pluralism of models of causal processes that may describe contributing factors in a given explanatory situation." (Mitchell 2003, 206) What might an integrative model look like for the debate over G×E?

4.2.2. The Call for an Integrative Solution

As it turns out, scientists investigating G×E are starting to ask this question about an integrative model themselves, attempting to move beyond the isolationist model advocated for so many years (Caspi and Moffitt 2006; Kendler 2005; Rutter 2006). The proposal by Avshalom Caspi and Terrie Moffitt, in particular, is worth looking at in detail, since it is their research on G×E that has lately received the most widespread attention from scientists and non-scientists alike (*Economist* 2002; Hamer 2002; Parens 2004). It was their research that I utilized to construct the reaction norm graph for the 5-HTT gene, stressful life events, and depression in Figure 17B earlier. With the increasing attention to their research on G×E, Caspi and Moffitt have recently recognized the usefulness of outlining for their readers the methodological and philosophical reasoning that guides them (Caspi and Moffitt 2006; Moffitt, Caspi and Rutter 2005). In the most recent of these offerings, Caspi and Moffitt (2006) develop a relationship between their statistical research on individual differences in depression and the scientists who study the underling mechanisms of depression. What, Caspi and Moffitt ask, is the relationship between the statisticians' work and that of the causal-mechanically-minded scientists? The answer, they argue, is that both are working together towards developing a *nomological network*. "A nomological network refers to the interlocking system of laws—the predicted pattern of

theoretical relationships—which define a construct." (Caspi and Moffitt 2006, 587) The idea is that the theoretical relationships under investigation by members of both the biometric and developmental research traditions may unite under this interlocking system of laws in order to integrate the fields. It is an idea borrowed from Lee Cronbach and Paul Meehl (1955), who introduced the concept of a nomological network in order to address the problem of validation in psychological tests.

There are several reasons, however, to be skeptical of how well the concept of a nomological network will capture the relationship between the two research traditions. For starters, appeals to timeless, universal generalizations (i.e., laws of nature) in biology have received extensive criticism from philosophers (Beatty 1995, 1997; Brandon 1997; Mitchell 1997, 2003; Sober 1997). It is argued that generalizations in biology are only contingently true because of their reliance on the contingencies of evolution (Beatty 1995).

In addition to this point, there is also a problem with how well the concept of a nomological network is designed to address the problems faced by biologists investigating the etiology of complex traits. As mentioned above, Cronbach and Meehl (1955) developed the idea to address problems with the validation of psychological tests. They were wrestling with a standard problem in psychology: What are psychological tests measuring? That is, to what extent do the constructs of intelligence or personality tests accurately map onto real features of individuals, such as general intelligence or temperament? But the problem with which Caspi and Moffitt (2006) are struggling pertains not to construct validity; it pertains to the main question of this essay—What is the relationship between the regular causal mechanisms responsible for individual development and the causes of variation responsible for individual differences? Thus,

we should expect a solution to Caspi and Moffitt's problem to reside in an account of this relationship rather than in the nomological network.

There is one more reason to be skeptical of how well the concept of a nomological network applies to the relationship between the statistical and the causal-mechanically-minded scientists. This has to do with the nomological network's reliance on (in addition to laws) *theories*. Like laws of nature, the prevalence of theories (in the standard linguistic sense) in biology has come under increasing scrutiny. For rather than searching for theories and theoretical relationships, biologists more often search for mechanisms as a means of providing biological explanations. And, in fact, when we turn to the work of Caspi and Moffitt along with the authors whom they cite as embodying the nomological network (Caspi and Moffitt 2006, 587), it is not the search for laws and theories that we find; it is the search for *mechanisms*:

> We hope that the present article will encourage further collaboration between genetic epidemiology and experimental neuroscience in a joint effort to unravel the complex mechanisms that underlie gene-environment interactions (Caspi and Moffitt 2006, 587).

> Identifying biological mechanisms through which genes lead to individual differences in emotional behavior is paramount to our understanding of how such differences confer risk for neuropsychiatric illness (Hariri and Holmes 2006, 182).

> In the Caspi et al. (2003) study, the impact of the serotonin polymorphism was seen only in the context of the environmental stressors. A mechanistic study of this nature is now

possible by, for example, pairing mice heterozygous for the serotonin transporter with a specific stress paradigm, such as maternal separation (Leonardo and Hen 2006, 132).

If the biometric and the developmental traditions cannot be integrated around their efforts to construct a nomological network, how can these traditions be integrated? The answer, I will now argue, comes from a shared problem—the elucidation of *difference mechanisms*.

4.3. Difference Mechanisms

Difference mechanisms are regular causal mechanisms made up of difference-making variables that take different values in the natural world. There is *regularity* in difference mechanisms; interventions made on variables in the mechanisms that change the values of the variables lead to different outcomes in the phenomena under investigation. There is also *variation* in difference mechanisms; interventions need not be undertaken to find differences in the outcomes because, with difference mechanisms, the variables already take different values in the natural world and so there already are differences in the outcomes. With difference mechanisms, *nature* is the interventionist.

The elucidation of difference mechanisms is a common denominator between the problem of elucidating the causes of variation responsible for individual differences and the problem of elucidating the regular causal mechanisms responsible for individual development. The result is an integrative relationship, not a reductive relationship. The relationship between the causes of variation responsible for individual differences and the regular causal mechanisms responsible for individual development is no more reductive than the relationship between 2/3 and 1/4. 2/3 cannot be reduced to 1/4, and 1/4 cannot be reduced to 2/3. The fractions can,

however, be integrated when a common denominator—12—is identified. Causes of variation responsible for individual differences cannot be reduced to regular causal mechanisms responsible for individual development, and regular causal mechanisms responsible for individual development cannot be reduced to causes of variation responsible for individual differences. They can, however, be integrated when a common denominator—difference mechanisms—is identified.

This section unpacks this metaphor, showing just how the causes of variation responsible for individual differences and the regular causal mechanisms responsible for individual development are related. In the next section I introduce the common philosophical idea of treating causes as difference-makers. Then, with that groundwork set, I expand on this notion by drawing attention to mechanisms where the difference-making variables naturally take different values. I will call these particular variables *difference-makers in development that take different values in the natural world*. It will be here that I provide the general framework for deriving the relationship between the various components of the biometric and the developmental research traditions. Finally, with that general framework in place, I return to the debate over G×E, arguing that G×E results from the *interdependence of difference-makers in development that take different values in the natural world*. Ultimately, then, the product will be a modified account of causal-mechanical explanation that captures both regularity and variation and which may be utilized to resolve the debates over G×E.

4.3.1. Causes as Difference-Makers

Philosophers have long conceived of causes as difference-makers. David Lewis, for instance, introduces his counterfactual theory of causation by explaining, “We think of a cause as

something that makes a difference, and the difference it makes must be a difference from what would have happened without it." (Lewis 1973, 557) Lewis, however, traces his idea to the second half of David Hume's definition: "…we may define a cause to be *an object, followed by another, and where all the objects, similar to the first, are followed by objects similar to the second*. Or in other words, *where, if the first object had not been, the second never had existed*." (Hume [1777] 1993, 51) The job for philosophers is explicating this idea of difference-maker; Lewis, for instance, attempts the explication with his theory of counterfactuals based on possible-world semantics.

A more recent explication of causes as difference-makers can be found in James Woodward's (2003) manipulationist conception of causation and causal explanation. The basic idea for Woodward is that scientists causally explain when they know how to manipulate. Manipulations are understood counterfactually. If some particular variable is a cause of some outcome, then manipulating the value of the variable would be a way of manipulating the outcome. These counterfactual experiments formulate and then answer, as Woodward explains, *what-if-things-had-been-different* questions; and, in so doing, they establish a pattern of counterfactual dependence between the explanandum (the thing to be explained) and the explanans (the thing or things that do the explaining).

Counterfactual dependence, for Woodward, is understood with the closely related concepts of intervention and invariance. An intervention consists of an idealized experimental manipulation of the value of some variable, thereby determining if it results in a change in the value of the outcome. So the counterfactuals are formulated in such a way that they show how the value of the outcome would change under the interventions that change the value of a variable; that is, they are formulated to show how the difference-makers make their difference.

Invariance, then, is a characterization of the relationship between variables (or a variable and an outcome) under interventions on Woodward's account. When there is an invariant relationship between a variable and an outcome, then that relationship is potentially exploitable for manipulation, and because of this it is a *causal* relationship.

Before applying Woodward's manipulationist conception of causation to the relationship between the causes of variation responsible for individual development and the causal mechanisms responsible for individual differences, I should say just why I utilize his version of causation as oppose to some other. For starters, the relata on Woodward's conception are *variables*. As we will see when we turn to an actual example from biology, variables in causal mechanisms are what scientists in the developmental and biometric research traditions investigate (rather than, say, *events*, the relata in Lewis's theory of causation). Also, Woodward's conception is specifically designed for capturing causal explanation in the *special sciences* (rather than, say, *physics*, the science examined by philosophers such as Salmon (1984, 1998) and Dowe (2000)). Designed for the special sciences, Woodward's development and employment of *invariance* also makes his conception well-suited for the developmental and biometric research traditions (rather than, say, *universal generalizations* or *laws of nature* employed by proponents of the deductive-nomological account of explanation such as Hempel and Oppenheim (1948)). Universal generalizations, as mentioned in the criticisms of the nomological network approach earlier, are unlikely to be found in the special sciences.

4.3.2. Difference-Makers in Development that Take Different Values in the Natural World

If we understand causes to be difference-makers, then a relationship between the regular causal mechanisms responsible for individual development and the causes of variation responsible for

individual differences becomes apparent. In short, the difference-making variables in the mechanisms simultaneously are the causes of variation when the difference-making variables take different values in the natural world. For example, the regular causal mechanisms responsible for the individual development of, say, depression consist of a number of variables (e.g., genes, neurotransmitters, brain systems, environmental insults), which may take different values in the natural world. Individual differences in depression result from the differences in the values of these difference-making variables in the mechanisms. Individual differences, then, are the effect of the difference-makers in individual development when the difference-makers naturally take different values.

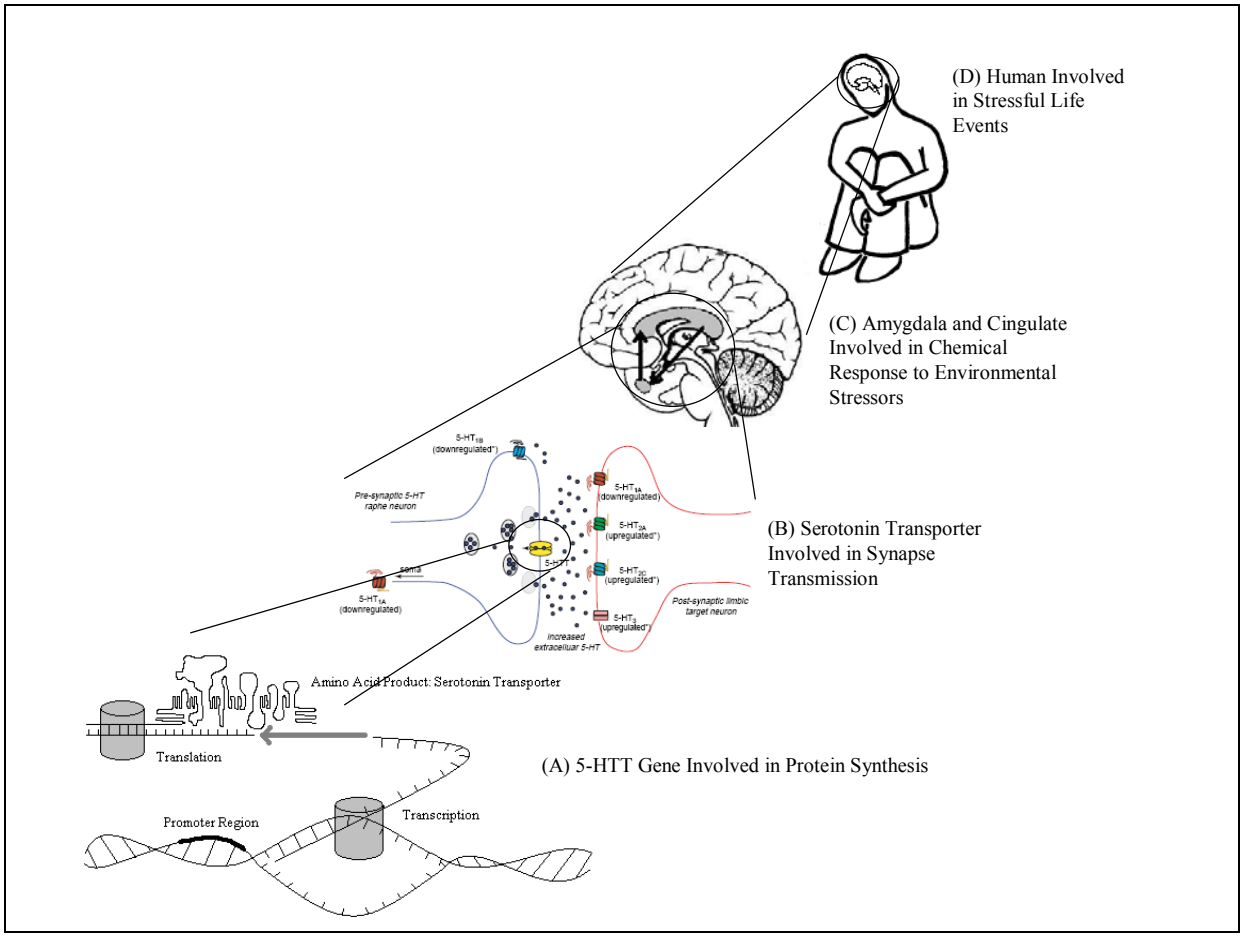

Figure 18. Mechanisms involved in the production of depression. (A) Protein synthesis at the molecular level. (B) Synapse transmission at the cellular level (modified from Hariri and Holmes 2006, Figure 1). Reprinted with the permission of TRENDS. (C) Amygdala-cingulate feedback at the brain-system level (modified from Bergmann 2000). Reprinted with the permission of The New Therapist at www.NewTherapist.com. (D) Experience of stressful life events at the organismal level.

To completely understand this idea, we should look at the example in more detail (Figure 18). The regular causal mechanisms responsible for depression are far from elucidated. But that need not prevent us from drawing on the portions that are known. The regular causal mechanisms responsible for depression are multi-level. At the (A) molecular level, the promoter region of the serotonin transporter gene (5-HTT) is involved in the mechanism of protein synthesis, which produces the serotonin transporter molecule as its amino acid product. At the (B) neural level, the serotonin transporter is involved in the mechanism of synapse transmission between pre- and post-synaptic neurons. At the (C) brain system level, the amygdala and the cingulate interact via feedback mechanisms that control chemical response to stressful environmental conditions. And at the (D) organismal level, humans experiencing environmental stressors such as, say, a looming dissertation defense are involved in mechanisms that generate stress. The 5-HTT gene, the serotonin transporter, the amygdala, the cingulate, and environmental insults, then, amount to several of the many difference-making variables in the regular causal mechanisms responsible for the individual development of depression. This difference-making capacity is determined because interventionist manipulations made on the values of these variables result in changes in the outcome. For example, inactivating the 5-HTT

gene in mice (a "knockout experiment") results in elevated levels of serotonin in certain regions of the brain, pointing to the role of the 5-HTT gene in mediating the reuptake of serotonin (Bengel et al. 1998).

I have intentionally modeled Figure 18 so as to resemble the earlier diagram (Figure 15) by Craver and Darden (2001) explicating the causal-mechanical explanation of spatial memory in rats. Whether it is spatial memory in rats or depression in humans, biologists provide explanations of complex behavioral traits by elucidating the regular causal mechanisms involved in the individual development of these traits. Interventions are made on variables to change the values of the variables so as to determine what the causal relationship is between a particular variable and another variable or the outcome of interest. Often these interventions force the variables to take unnatural values. For example, a gene knockout experiment attempts to determine the causal relationship between a genetic variable and a phenotypic trait by forcing the genetic variable to take an unnatural value—non-existence. Likewise, an animal deprivation study attempts to determine the causal relationship between an environmental variable and a phenotypic trait by forcing the environmental variable to take an unnatural value—extreme deprivation. Scientists, by isolating a variable and manipulating its value to unnatural extremes, try to hold all other relevant variables constant and then attribute the change in outcome to the change in the variable's value, thereby identifying the causal relationship between the variable and the outcome. The ideal explanation of depression, then, would amount to an identification of all the difference-making variables as well as an account of how those difference-makers make their difference in the regular causal mechanisms responsible for the individual development of depression.

But what about the other question: How do some people come to experience a complex trait such as depression, while others do not? This is the question that demands an answer about *variation*. In providing a causal explanation for this question, we cannot just point to how the difference-making variables make their difference in the regular causal mechanisms responsible for the individual development of depression by, say, knocking out the 5-HTT gene, for this does not yet address the question about why there are actual individual *differences* in depression. This is because we do not expect to find actual individuals with this extreme value taken by this variable in the natural world. To address individual differences, then, we must determine what values the variables take in the natural world and determine how the differences in those real-world values lead to actual individual differences in the outcome. That is, we must identify the *difference-makers in development that take different values in the natural world*.

Waters (Forthcoming) draws attention to a similar concept, albeit in the confines of different debates. Waters' debates of interest also reside in biology and the philosophy of biology; he is concerned with disputes over the relationship between classical, Mendelian genetics and molecular genetics, as well as with disputes over causal parity between genes and other variables involved in the causal mechanisms responsible for the development of traits. While any variable involved in the causal mechanisms responsible for the development of a particular trait may be a *potential* difference-maker, Waters argues that only certain variables are *actual* difference-makers because only certain variables actually take different values in the natural world.

Biometric behavioral geneticists take it upon themselves to answer the question about the causes of variation responsible for individual differences. They have identified the different forms of the promoter region in the 5-HTT gene (s/s, s/l, and l/l) and determined that a small

portion of total variation in depression (3-4%) is attributable to individual differences in the value of this genetic variable (Lesch et al. 1996). More proximally, behavioral geneticists have found differences in the value of the 5-HTT variable to be related to differences in amygdala activity; individuals with the "s" allele exhibit much greater amygdala activity than l/l individuals when presented with images of fearful and angry human faces (Hariri et al. 2002). Also, individuals with the "s" allele, compared with l/l individuals, exhibit a weaker coupling between the amygdala and the cingulate in the feedback mechanisms responsible for the chemical response to environmental stressors (Pezawas et al. 2005).

The biometric behavioral geneticists seek causes of variation responsible for individual differences in depression. But these causes of variation responsible for individual differences are not distinct from the difference-making variables in the regular causal mechanisms responsible for individual development. *They are not isolated at different levels of analysis*. Rather, the difference-making variables in the regular causal mechanisms simultaneously are the causes of variation just when the difference-making variables take different values in the natural world, or, in Waters' locution, when the variables are *actual* difference-makers. The promoter region of the 5-HTT gene, for example, is a difference-maker in the individual development of depression, as the knockout experiments reveal. Though the precise mechanism has not been elucidated, the variable seems to make a difference in how the amygdala coordinates with the cingulate to respond to environmental stressors. The promoter region of the 5-HTT gene is simultaneously a cause of variation because it takes different values in the natural world (s/s, s/l, and l/l), and the differences in the value of the genetic variable are responsible for slight individual differences in depression. Individual differences, then, are the effect of difference-makers in individual development when the difference-makers naturally take different values.

Difference mechanisms, remember, are regular causal mechanisms made up of difference-making variables that take different values in the natural world. In the case of depression, the promoter region of the 5-HTT gene is one of these variables that naturally takes different values; it is a difference-maker in development that takes different values in the natural world. Consider Figure 19: At the molecular level, *differences* in the promoter region of the 5-HTT gene are responsible for *differences* in the number of serotonin transporters that are available for synapse transmission at the cellular level; the shorter promoter region generates relatively less serotonin transporter molecules. *Differences* in synapse transmission, then, are responsible for *differences* in amygdala-cingulate feedback at the brain-system level; individuals with an "s" allele have a relatively weaker coupling between the amygdala and the cingulate in response to environmental stressors, leading to relatively prolonged exposure to negative emotional states. And *differences* in amygdala-cingulate feedback are responsible for *differences* in depression; relatively prolonged exposure to negative emotional states puts one at risk of developing depression if environmental stressors are encountered often enough.

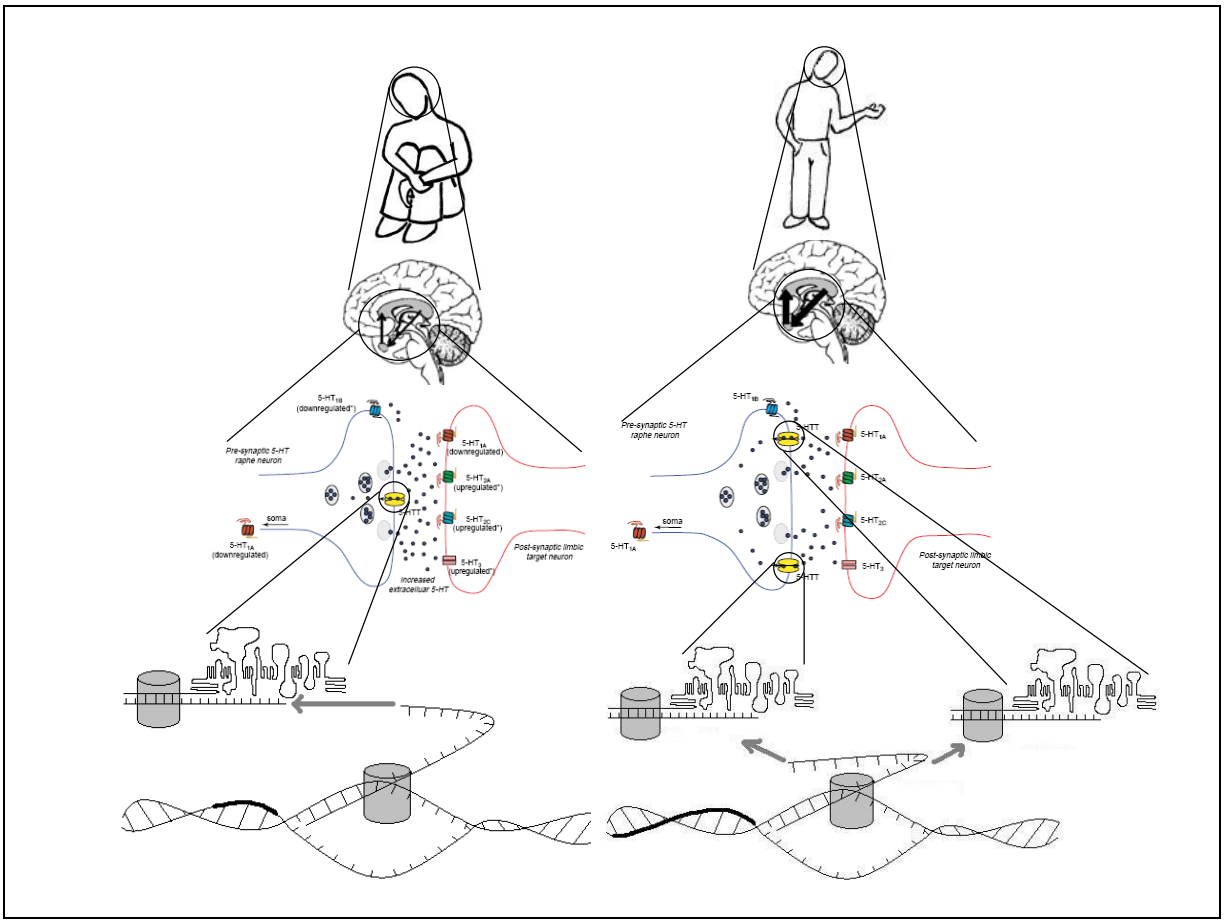

Figure 19. Difference mechanisms involved in the production of depression.

Philosophers of science interested in the concept of a mechanism have focused thus far on how causal explanation arises from the elucidation of the regular causal mechanisms responsible for the individual development of a particular trait. They have focused thus far on how scientists provide causal explanations by identifying and manipulating the difference-making variables in regular causal mechanisms in an attempt to determine how those difference-makers make their difference. I have tried to show in this section, however, that the scientists in the biometric tradition are involved in a slightly different enterprise, and so the philosophy of mechanisms must be slightly revised if it is to be extended to cover the causal explanations of individual differences sought by scientists in the biometric tradition. Focused on individual

differences, these scientists do not provide causal explanations by showing how difference-makers make their difference in the regular causal mechanisms. Rather, they causally explain by showing how or to what extent *differences* in the natural values of the difference-makers result in individual *differences* in the trait or in individual *differences* in the values of more proximal variables. I have also tried to show, however, that a scientist's focus on differences does *not* somehow isolate her at a unique level of analysis. Scientists in the biometric tradition are examining causes of variation responsible for individual differences in a population, but these causes of variation just are the difference-makers in the regular causal mechanisms when the difference-makers take different values in the natural world. And the individual differences just are the effect of the difference-makers in development that naturally take different values.

Of course, scientists in the biometric tradition and scientists in the developmental tradition do elucidate difference mechanisms with different approaches. Scientists in the biometric tradition ask *how-much?* questions about the causes of variation and utilize statistical methodologies to answer these questions. Scientists in the developmental tradition ask *how?* questions about the regular causal mechanisms and employ interventionist methodologies to answer these questions. But, again, these differences in question and methodology do not isolate the traditions at different levels of analysis. They are just different approaches to elucidating the difference-makers in development that take different values in the natural world, which are responsible for difference mechanisms. Causal-mechanically-minded scientists attempt to determine *what* the difference-makers are in the regular causal mechanisms responsible for the individual development of a particular trait and determine *how* those difference-makers make their difference within the mechanism. Interventionist methodologies are employed to artificially change the values of the difference-making variables so as to make this determination.

Biometrically-oriented scientists move beyond this artificial intervention and determine *which* of the difference-makers naturally take different values and *how much* of the difference in outcome can be attributed to differences in the values that a difference-making variable takes. Statistical methodologies are employed to study populations with differences in the outcome and differences in the values of a particular variable so as to make this determination.

Ultimately, then, there are very real differences between the biometric and the developmental research traditions. Scientists in the biometric tradition employ statistical methodologies to answer how-much? questions about the causes of variation responsible for individual differences. Scientists in the developmental tradition employ interventionist methodologies to answer how? questions about the regular causal mechanisms responsible for individual development. But scientists in these two traditions are not isolated at different levels of analysis. The twain shall meet. They meet at a common denominator—difference mechanisms.

Before returning to the disputes over G×E, I should highlight the fact that difference mechanisms are a common denominator for more than just the biometric and developmental research traditions. In joining regularity and variation, difference mechanisms integrate any discipline(s) that investigate these two features of the biological world. Mayr's (1961, 1982) influential distinction between *proximate* and *ultimate* causes provided a causal framework on which to situate the distinction between different levels of analysis discussed in section 4.2.1 above. On Mayr's account, the proximate causes of anatomy and physiology are investigated by functional biologists, while the ultimate causes of phylogeny are investigated by evolutionary biologists. Many scientific investigations, however, bridge this divide. This dissertation has been devoted to examining one such example: investigations into the etiology of complex traits. It is

by no means alone. Evolutionary-developmental (or evo-devo) biologists examine, for example, the ways in which developmental mechanics constrain evolutionary possibilities. That is, evolutionary-developmental biologists examine how only certain difference mechanisms allow for viable individual development and then generate the variation upon which natural selection can act. Difference mechanisms, then, are a common denominator between the biometric and developmental research traditions. But more generally, difference mechanisms are a common denominator between proximate and ultimate biology.

4.3.3. The Interdependence of Difference-Makers in Development that Take Different Values in the Natural World

Appeals to different levels of analysis in the nature-nurture debate draw a firm line between a legitimate, statistical notion of interaction and a muddle-headed, developmental notion of interactionism. It is certainly the case that G×E does not simply follow from each and every interaction between genes and the environment during individual development, and so it is misleading to point to developmental interactions and assume that G×E must follow. In this sense, the following criticism of the biometric tradition *is* muddle-headed: "An analysis of variance abstracts from (actual) interaction effects and thus cannot offer an accurate picture of development…. Behavioral geneticists, then, should acknowledge that an analysis of variance is a statistical method that does not fit reality and should be judged against the background of the best material model we have of development, which is one of dynamics and interactions." (Vreeke 2005, 44) Perhaps the ignominious "interactionism" should be reserved for such confusions.

But must all invocations of individual development in considerations of G×E be of the muddle-headed sort? Not at all. As I argued in the last section, the difference-making variables in the regular causal mechanisms simultaneously are the cause of variation when the variables take different values in the natural world, and individual differences are the effect of difference-makers in development. Now, if difference-making variables interact during the course of individual development, then a variable (X) taking a particular value (X_1) in interaction with another variable (Y) will likely lead to different outcomes depending on the value taken by Y (Y_1 vs. Y_2); but this difference in outcomes may have been quite different if variable X had taken a different value (X_2). Or, to continue with the depression example, the regular causal mechanisms responsible for the individual development of depression consist of a number of variables (the 5-HTT gene, the serotonin transporter, the amygdala, the cingulate, environmental insults, etc.), some of which take different values in the natural world. As explained in the last section, individual differences in depression result from differences in the value of the promoter region of the 5-HTT gene and from differences in the value of the number of stressful life events experienced. But, as we will see, individual differences in depression *also* result from differences in particular combinations of 5-HTT and stressful life events as these two variables have the opportunity to interact during the course of individual development.

Again, we should look at this example in more detail to completely understand the idea. As mentioned in the last section, we know that the promoter region of the 5-HTT gene is a difference-making variable in the regular causal mechanisms responsible for the individual development of depression; we also know that this variable takes different values in the natural world; and we also know that individual differences in depression result from the different values of this difference-maker in development, although the variation attributable to this difference-

maker is very small (3-4%). We know that stressful live events are a difference-making variable in the regular causal mechanisms responsible for the individual development of depression; we also know that this variable takes different values in the natural world; and we also know that individual differences in depression result from the different values of this difference-maker in development.

Now, suppose that no matter how many stressful life events were experienced, having the s/s value of the promoter region would increase the probability of individuals experiencing depression by 10% relative to having the s/l value, and by 20% relative to having the l/l value. Likewise, suppose that no matter which value of the promoter region was had, every stressful life event experienced increased the probability of individuals experiencing depression by 5%. I have been saying "suppose" because the hypothetical situation I am describing here was depicted in the hypothetical reaction norms graphed in Figure 17(A) and is recreated in Figure 20(A) with the individuals from the discussion of difference mechanisms now mapped onto the reaction norms. This case, remember, arises when the genetic and environmental sources of variation are additive and the reaction norms are parallel. The total phenotypic variation in depression is fully accounted for by pointing to the separate differences in the value of the genetic variable and the differences in the value of the environmental variable. But remember this hypothetical situation was first introduced only to contrast it with the empirical data from the *actual* reaction norms graphed in Figure 17(B). In the empirical example, as the reaction norms clearly show, there is variation due to G×E in addition to the variation resulting from differences in the value of the genetic variable and differences in the value of the environmental variable. Thus, there is *no way* to predict who is more at risk of developing depression—either individuals with the s/s, s/l, or l/l genotype—before the environmental conditions of development are experienced (see Figure

20(B)). Caspi and Moffitt found that in environments with multiple stressful life events experienced, individuals with the s/s genotype are at *greater* risk of developing major depression; while, in environments without multiple stressful life events experienced, individuals with the s/s genotype are at *less* risk of developing major depression. That is, there is a *change in rank* for the relationship between the genetic variable, the environmental variable, and the phenotypic trait (Lynch and Walsh 1997).

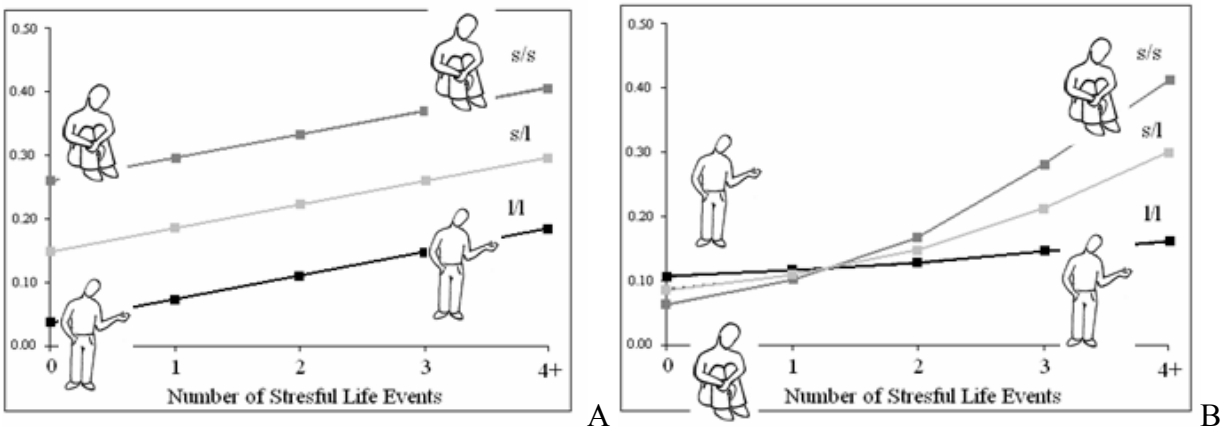

Figure 20. Reaction norm graphs for probability of depression (y-axis), promoter region of serotonin transporter gene (s/s vs. s/l vs. l/l), and number of stressful life events experienced (x-axis). (A) Hypothetically parallel reaction norms. (B) Reaction norms drawn from empirical data with *change in rank* highlighted.

With the interaction in this empirical example in mind, let us return to the debate between the defenders of the biometric tradition and their causal-mechanically-minded critics who appeal to the complications posed by individual development. What precisely does the concept of G×E *mean*? Does it incorporate an element of individual development, or not? The suggestion on offer by those who appeal to different levels of analysis is that G×E *just is* the breakdown in

additivity between main effects measured by ANOVA. But we can go on and ask, what *causes* this breakdown in additivity? The answer is that this breakdown in additivity is caused by the *interdependence of difference-makers in development that take different values in the natural world*. That is, G×E results from differences in particular combinations of genetic and environmental variables when both variables are difference-makers in development that naturally take different values and the difference that each variable makes is itself dependent upon the difference made by the other variable. Difference mechanisms, then, arise when difference-making variables in the regular causal mechanisms take different values in the natural world. And difference mechanisms also arise when there are differences in the interactions between difference-making variables depending on which values the difference-makers take.

4.4. Conclusion

Philosophers of science, in recent years, have developed a renewed interest in mechanisms. The account of causal-mechanical explanation on offer is designed to capture causal explanations of *regularity*. For example, the account is designed to show how a causal explanation follows from the elucidation of the regular causal mechanisms responsible for the individual development of a complex behavioral trait such as spatial memory in rats or depression in humans. Focused on regularity, however, this account of causal-mechanical explanation neglects *variation*, one of the core features of the biological world and, in turn, of biological explanation. That is, the account fails to show what role in causal explanation is played by the elucidation of the causes of variation responsible for individual differences in spatial memory or depression.

The purpose of this chapter was to modify the mechanical program so as to capture both regularity *and* variation. The task was to explicate the relationship between the regular causal

mechanisms responsible for individual development and the causes of variation responsible for individual differences. This relationship, as it turns out, is not just of interest to the armchair philosopher. For it is precisely this relationship that resides at the heart of debates over G×E between scientists in the biometric research tradition and scientists in the developmental research tradition, which was detailed in Chapters 2 and 3.

The solution was the identification of a shared problem, or a "common denominator," between the biometric and developmental research traditions—the elucidation of what I called *difference mechanisms*. Difference mechanisms are regular causal mechanisms made up of difference-making variables that take different values in the natural world. With this general framework, the relationship between the regular causal mechanisms responsible for individual development and the causes of variation responsible for individual differences becomes apparent. The difference-making variables in the regular causal mechanisms responsible for individual development simultaneously are the causes of variation responsible for individual differences when the variables take different values in the natural world; that is, when the variables are *difference-makers in development that take different values in the natural world*. And individual differences are the *effect* of these difference-makers in development that naturally take different values. This general framework was then applied to the debates over G×E. G×E, in contrast to appeals to different levels of analysis, does incorporate an element of individual development; G×E results from the *interdependence of difference-makers in development that take different values in the natural world*. Ultimately, then, the product was a modified account of causal-mechanical explanation that captured both regularity and variation, and which was utilized to resolve the debates over G×E.

5. RATS! SO WHAT *IS* G×E?

Abstract. Three concepts of genotype-environment interaction, or G×E, have now been defined: a biometric concept ($G\times E_B$), a developmental concept ($G\times E_D$), and what may be called an interdependent-difference-makers concept. So what *is* G×E? Or, more specifically, what is the relationship between these three concepts? The thesis of this chapter will be the following: Following from the integrative framework developed in the last chapter, $G\times E_B$ and $G\times E_D$ may be integrated under the interdependent-difference-makers concept of G×E. More specifically, G×E results from differences in unique, developmental combinations of genotype and environment when both variables are difference-makers in development that take different values in the natural world and the difference that each variable makes is itself dependent upon the difference made by the other variable; a breakdown in additivity between main effects is a measure of this interdependence of difference-makers that naturally take different values. More succinctly: the interdependent-difference-makers concept of G×E is just a general, causal-mechanical interpretation of $G\times E_D$, of which $G\times E_B$ is a statistical measure.

5.1. Introduction

Genotype-environment interaction, or G×E, is the result of the breakdown in additivity between genotypic and environmental sources of variation, which is measured by a statistical methodology such as the analysis of variance. G×E is the result of differences in unique, developmental combinations of genotype and environment. G×E is the result of the interdependence of difference-makers in development that take different values in the natural world. At this point, I have now defined three different concepts of G×E: a biometric concept ($G\times E_B$), a developmental concept ($G\times E_D$), and what may be called an interdependent-difference-makers concept. So what *is* G×E? Or, more specifically, what is the relationship between these three concepts?

In Chapters 2 and 3 I teased apart the separate components of the biometric and developmental research traditions so as to identify the different axes upon which the debates over G×E were disputed. The goal of those chapters was also to show how two different concepts of G×E—$G\times E_B$ and $G\times E_D$—were situated in these separate traditions. The concepts originated in

the work of Fisher and Hogben, persisted through mid-twentieth century population and developmental genetics, and on into the IQ Controversy of the 1970's. In Chapter 4 I showed how the two research traditions could be integrated via a shared problem—the elucidation of difference mechanisms. But there was no mention of either $G\times E_B$ or $G\times E_D$ in that discussion of integration. The purpose of this chapter is to pull the separate threads of these earlier chapters together. That is, the goal is to utilize the integrative model developed in the last chapter and apply it to the divide between $G\times E_B$ and $G\times E_D$. The thesis, in short, is that $G\times E_B$ and $G\times E_D$ may be integrated under the interdependent-difference-makers concept of G×E. More specifically, G×E results from differences in unique, developmental combinations of genotype and environment when both variables are difference-makers in development that take different values in the natural world and the difference that each variable makes is itself dependent upon the difference made by the other variable; a breakdown in additivity between main effects is a measure of this interdependence of difference-makers. Another way to think about the relationship is to see that the interdependent-difference-makers concept of G×E is just a more general, causal-mechanical reinterpretation of $G\times E_D$, of which $G\times E_B$ is a statistical measure.

The best way to unpack this idea will be to look at an example in some detail. Although the research by Avshalom Caspi, Terrie Moffitt, and their colleagues has received the greatest attention in recent years (Caspi et al. 2002; Caspi et al. 2003; Caspi et al. 2005), no empirical result of G×E has received more attention in the history of the nature-nurture debate than the study by Roderick Cooper and John Zubek on different strains of rats raised in different environments (Cooper and Zubek 1958). This study was only mentioned in passing in the last chapter, but it will be worth looking at in detail in an effort to reveal the relationship between the various concepts of G×E. In the next section I will highlight the key features of the separate

legacies of $G \times E_B$ and $G \times E_D$, since it has been a chapter since the reader has had to think about the distinction between the two concepts. The goal here will be to remind ourselves of the *divide* between $G \times E_B$ and $G \times E_D$ that emerged over the years. I will then bridge that divide in section 5.3, utilizing the interdependent-difference-makers concept to build that bridge and applying it to the Cooper-Zubek study.

5.2. $G \times E_B$ vs. $G \times E_D$

The purpose of this section will be to draw out the highlights of the separate legacies of the biometric and the developmental concepts of G×E. The legacies were traced in Chapters 2 and 3 from their origin*(s)* to the IQ Controversy. But it will be worth consolidating that history in one place and also tracing those legacies into the present so as to convey the fact that the divide between $G \times E_B$ and $G \times E_D$ persists even today.

5.2.1. $G \times E_B$

The biometric concept of G×E originated in the work of R. A. Fisher. Fisher, while at the Rothamsted Agricultural Research Station, developed many of the now-standard statistical methodologies designed to measure the relative contributions of nature and nurture to individual differences in a population, such as the analysis of variance (ANOVA), the design of experiments, and the statistical significance test. While developing these tools, Fisher quickly realized that the presence of what he called "non-linear interactions between heredity and environment" posed a potential complication for the partitioning of causes of variation. If such interactions existed, he realized, the main effects of genotype and environment did not add up to the total phenotypic variance, and inferences about what different genotypic groups would do in

different environments became highly suspect. Fisher, in order to consider the empirical reality of this potential complication, undertook an examination of different potato varieties grown in different fertilizers (Fisher and Mackenzie 1923). In an early test of statistical significance, however, Fisher found the interaction effect, or the "deviations from summation formula" as he called it, to be no greater than chance, leading him to conclude that there was no evidence of interaction. This result was evidently quite conclusive for Fisher; he mentioned the possible complications raised by the "interactions of causes" in the chapter on ANOVA in his *Statistical Methods for Research Workers* (Fisher 1925), but he again referenced his 1923 study with Mackenzie and concluded that there was no evidence of interaction. Fisher, thus, took cases of statistical non-additivity to be rare in nature. And even when a case did arise, Fisher simply encouraged a transformation of scale to make the statistical complication go away.

In summary, Fisher operated in the biometric research tradition (Table 9). The main problem on which Fisher was focused was the partitioning of the *relative contributions of nature and nurture responsible for individual differences* in populations. His approach to causation involved an investigation into the *causes of variation* responsible for these individual differences. He asked, *how much* of the variation in a particular population was due to individual differences in heredity or environment? And he sought to answer those questions with his population-level, *statistical methodologies*. Fisher's concept of interaction was developed in this biometric tradition. His resulting *biometric* concept of G×E, or $G \times E_B$, may be defined as a statistical measure of the breakdown in additivity between genotypic and environmental sources of variation, which is generated by a statistical methodology such as the analysis of variance.

Components	**Biometric Research Tradition**
Problem	Individual Differences
Approach to Causation	Causes of Variation
Causal Question	How Much?
Methodology	Statistical
Concept of Interaction	Biometric—$G \times E_B$

Table 9. The components of the biometric research tradition.

The biometric research tradition in biology did not end with Fisher. It was carried into the mid-twentieth century in the form of population genetics (Provine 2001). Jay Lush, one of the leaders in American agricultural genetics, continued on this tradition as well as the biometric interpretation of interaction. In his influential *Animal Breeding Plans* (1937), Lush echoed Fisher's conclusions about G×E: Cases were rare; and even if cases did arise, they could be statistically eliminated with a transformation of scale. Indeed, G×E was worth little more than a dismissive footnote for Lush: "For some extreme examples of nonaddtive combination effects of heredity and environment consult chapter 5 of Hogben's *Nature and Nurture*." (Lush 1937, 64, fn.)

In 1960, quantitative behavioral genetics emerged as a discipline out of population genetics (Griffiths and Tabery Forthcoming). Quantitative behavioral geneticists borrowed the statistical tools of population genetics and applied them to complex behavioral traits, such as IQ in humans. In the 1970's, Arthur Jensen carried on the biometric research tradition, applying the biometric tools to the question of racial differences in IQ scores (Jensen 1969). Jensen, operating in the biometric tradition, was interested in the *causes of variation* responsible for *individual differences* in IQ; he asked *how-much?* questions about these causes of variation and employed *statistical methodologies* to answer the questions. And Jensen echoed Lush and Fisher's

conclusions about G×E: Cases were rare; and even if cases did arise, they could be statistically eliminated with a transformation of scale. Jensen can also be credited with introducing the distinction between statistical interaction and muddle-headed "interactionism" discussed in the last chapter. Jensen, when criticized for overlooking the complications posed by developmental interactions between genotype and environment, responded, "'interactionism' has become merely a substitute for extreme environmentalism. … This position has arisen from a failure to understand the real meaning of the term 'interaction' as it is used in population genetics; but even more it is the result of failure to distinguish between (a) the *development* of the individual organism, on the one hand, and (b) *differences* among individuals in the population." (Jensen 1973, 49)

The legacy of $G \times E_B$ can be traced into contemporary philosophy of science. In his *Making Sense of Heritability* (2005), Neven Sesardic offered a defense of the biometric research tradition. With regards to G×E, Sesardic echoed Fisher and Lush and Jensen: Cases are rare; and even when cases do arise, they can be statistically eliminated with a transformation of scale (ibid, 68-70). And, lest philosophers mistakenly invoke individual development in discussions of individual differences, Sesardic also offered up his own distinction between *statistical* interaction (interaction_s) and *commonsense* interaction (interaction_c). Too often, Sesardic complained, philosophers of science mistakenly assumed that developmental interactions between genotype and environment were related to population-level interactions between sources of variation, but this only confused levels of analysis (ibid, 73).

5.2.2. $G \times E_D$

The developmental concept of G×E originated in the work of Lancelot Hogben. Hogben, from very early in his career, took an interest in the causal-mechanics of individual development. While acting as Chair of Social Biology at the London School of Economics, Hogben turned his attention to the British eugenics movement, which had reached its influential climax in the 1930's. At the LSE, Hogben utilized his appreciation for the causal-mechanics of development to attack the attempts made by eugenicists to separate the contributions of nature and nurture to complex human traits such as pauperism, alcoholism, or feeblemindedness. G×E figured prominently in this attack. In 1932, Hogben published his *Genetic Principles in Medicine and Social Science*, where he drew attention to a *third class of variability* in addition to hereditary variation and environmental variation: "that which arises from the combination of a particular hereditary constitution with a particular kind of environment." (Hogben 1932, 98) At that time, though, Hogben offered his readers only a qualitative example of this third class: the abnormal abdomen strain of *Drosophila*, which resembled the wild type when grown in dry environments but which developed abdominal mutations when grown in humid environments. In 1933, Hogben revised this point with a quantitative example, a necessity if he was to engage the statistical methods employed by the eugenicists. Hogben, in correspondence with Fisher, introduced Krafka's (1920) data on different *Drosophila* strains, environmental variation in temperature, and phenotypic variation in number of eye facets. Hogben took this case to be an example of the "intimate sense in which differences of genetic constitution are related to the external situation in the process of development."[58] This example then anchored the portion of his William Withering Memorial Lectures devoted to the "Interdependence of Nature and Nurture" (Hogben 1933). Hogben, thus, took cases of interaction to reflect the developmental relationship between

[58] Hogben to Fisher, 23 February 1933, R. A. Fisher Papers (Series I, Hogben, L.), University of Adelaide.

genotype and environment. When empirical cases were identified, they were not to be eliminated as a nuisance; they were to be embraced as bearing important information about development and variation.

In summary, Hogben operated in the developmental research tradition (Table 10). The problem on which Hogben was focused was unraveling the way in which variation in a population arose from the relationship between genotype and environment during *individual development*. His focus was on the *causal mechanisms* of individual development. He asked, *how* do differences in genotype and differences in environment relate during individual development to generate differences in phenotype? And he employed or sought out *interventionist methodologies*, such as those undertaken by Krafka, to manipulate these variables and monitor the phenotypic outcomes. Hogben's concept of interaction was developed in this developmental tradition. His resulting *developmental* concept of G×E, or $G \times E_D$, was his third class of variability. It may be defined as variation that results from differences in unique, developmental combinations of genotype and environment.

Components	**Developmental Research Tradition**
Problem	Individual Development
Approach to Causation	Causal Mechanisms
Causal Question	How?
Methodology	Interventionist
Concept of Interaction	Developmental—$G \times E_D$

Table 10. The components of the developmental research tradition.

The developmental research tradition in biology did not end with Hogben. It was carried into the mid-twentieth century in the form of developmental genetics. Conrad Hal Waddington,

one of the leaders in British developmental genetics, continued on this tradition as well as the developmental interpretation of interaction. In his influential *The Strategy of the Gene* (1957), Waddington echoed Hogben's conclusions about G×E: They were developmental in nature, and they were of utmost importance for understanding development and variation. "Now from the point of view of the theory of evolution," Waddington explained, "such special interactions between genotypes and environments are obviously by no means negligible. In fact, the whole of adaptive radiation, including the formation of local races, turns on the way in which particular genotypes fit into certain environments; that is to say, on this very factor of genotype-environment interaction." (ibid, 100)

In the 1970's, Richard Lewontin and David Layzer attacked Arthur Jensen's discussion of the causes of variation responsible for individual differences in IQ. The real purpose of genetics, they argued, was to elucidate the *causal mechanisms* responsible for *individual development*, to answer *how?* questions about these causal mechanisms with *interventionist methodologies*. And Lewontin and Layzer echoed Waddington and Hogben's conclusions about G×E: They were developmental in nature, and they were of utmost importance for understanding development and variation (Layzer 1974; Lewontin 1974).

The legacy of $G \times E_D$ can be traced into contemporary philosophy of science. Lewontin and Layzer's discussions G×E and the complications it poses to heritability estimates have left a lasting impression on the philosophy of biology. As Stephen Downes recently summarized, "The point of departure for many philosophers criticizing heritability analysis is Lewontin's (1974) paper on the analysis of variance. … The current consensus among philosophers of biology is that heritability analyses are misleading about the genetic causes of human traits." (Downes 2004, section 3)

5.3. $G \times E_B$ and $G \times E_D$

On the one hand, scientists in the biometric tradition define G×E in statistical terms—in terms of a breakdown in additivity measured by a statistical methodology such as ANOVA. On the other hand, scientists in the developmental tradition define G×E in developmental terms—in terms of developmental relationships between genotype and environment. The biometricians criticize the developmentalists for confusing levels of analysis. And the developmentalists criticize the biometricians for ignoring development. Must we decide between the biometricians' $G \times E_B$ and the developmentalists' $G \times E_D$? No. The two may be integrated via the interdependent-difference-makers concept of G×E that I introduced in the last chapter. G×E results from differences in unique, developmental combinations of genotype and environment when both variables are difference-makers in development that take different values in the natural world and the difference that each variable makes is itself dependent upon the difference made by the other variable; a breakdown in additivity between main effects is a measure of this interdependence of difference-makers. Let me explain precisely what I mean by this relationship, starting with a more detailed examination of $G \times E_B$.

5.3.1. What Is an Analysis of Variance?

The place to best begin understanding $G \times E_B$ is by examining just how it emerges out of the breakdown in additivity measured by a statistical methodology such as ANOVA. Fisher's statistical methodology is now one of the standard resources in any biometricians' toolbox. In textbooks, such as Robert R. Sokal and F. James Rohlf's *Biometry*, ANOVA is described in grand terms: "Once it is understood, analysis of variance is a tool that can provide an insight into

the nature of variation of natural events, into Nature in short, which is possible of even greater value than the knowledge of the method as such." (Sokal and Rohlf 1995, 179). *But what is ANOVA?*[59] More directly, how does ANOVA provide insights into the nature of variation, or into the nature of causes of variation?

5.3.1.1. Mill's Methods

John Stuart Mill's discussion of the methods of experimental inquiry provides the foundation for any study of methodological analyses into causation (Mill 1974 [1843]). In his *System of Logic, Ratiocinative, and Inductive*, Mill introduced what have come to be called "Mill's Methods," an exposition of the various ways in which causes are identified in nature. The first two methods are the simplest: the *method of agreement* and the *method of difference*. If the goal is to identify the cause (A) of some effect (*a*), the method of agreement works by finding cases which agree in one circumstance (A) but differ in every other and yet still produce the effect *a*. In other words, if A, B, C, D, and E are various causes, and *a*, *b*, *c*, *d*, and *e* are various effects, the cause of *a* can be attributed to A when the following two cases are compared:

$$A\ B\ C \rightarrow a\ b\ c$$

$$A\ D\ E \rightarrow a\ d\ e$$

The cases hold only cause A in agreement and then hold only effect *a* in agreement as well. The method of difference works by finding cases resembling each other in every respect *except* the effect under consideration and identifying the cause that is missing. For instance, the cause of *a* can be attributed to A when the following two cases are compared:

$$A\ B\ C \rightarrow a\ b\ c$$

[59] I am not the first to ask this question. The question is addressed at length by Speed (1987). However, Speed is concerned with the *mathematical* nature of ANOVA. In contrast, I am concerned with the *experimental* nature of ANOVA, in how it provides insights into causation and variation.

$$B\ C \rightarrow b\ c$$

The cases differ only in having or lacking cause A and then differ only in having or lacking effect *a*. The method of agreement and the method of difference share an important commonality: both are methods of *elimination*. That is, the method of agreement is based on the principle that whatever *can* be eliminated is *not* connected to the effect. And the method of difference is based on the principle that whatever *cannot* be eliminated *is* connected to the effect.

Of course, Mill readily recognized the fact that nature did not always lend itself to such simple inquiries. An investigator, for example, cannot always eliminate a cause. Mill pointed out that *permanent causes* cannot be eliminated, making both methods of agreement and difference inapplicable (ibid, 398). For instance, if an investigator is interested in the causal relationship between the tides and the moon, the investigator cannot eliminate the moon so as to examine what effect is had on the tides by its absence. What to do? The solution, Mill offered, was the *method of concomitant variation*. "Though we cannot exclude an antecedent altogether," Mill admitted, "we may be able to produce, or nature may produce for us, some modification in it." (ibid, 400) The method of concomitant variation works by comparing cases which are identical except for a variation in one cause A which leads to a subsequent variation in the effect *a*:

$$A_1\ B_1\ C_1 \rightarrow a_1\ b_1\ c_1$$

$$A_2\ B_1\ C_1 \rightarrow a_2\ b_1\ c_1$$

Though the causes cannot be eliminated so as to apply either the method of agreement or the method of difference, the causal relationship can be investigated by comparing the variations in the cause with the variations in the effect. For instance, variations in the moon's position can be compared to variations in the tides. In addition to the problem of permanent causes, Mill also recognized the fact that causes and effects are rarely in a simple, one-to-one relationship with

each other as is assumed in the A → *a* depictions above. Causal relationships are usually much more complex in nature. More specifically, there may be a *plurality of causes* responsible for any given effect (ibid, Book III, chapter 10). In such situations, the causes must somehow be isolated so as to examine their effects apart from the other causes.[60]

5.3.1.2. Fisher's Methods

With Mill's methods now introduced, we may now turn to Fisher's methods. ANOVA, or more specifically a two-way ANOVA, is in its essence an application of the method of concomitant variation to a situation involving a plurality of causes.[61] Consider a hypothetical population of rats. We are interested in what causal role genes and the environment play in a phenotypic trait such as spatial memory. As I discussed in the last chapter, spatial memory in rats can be operationalized by exposing rats to a Hebb-Williams maze test (Hebb and Williams 1946). Rats start at one corner of a maze and attempt to navigate to a food source at the other end. The rats, after attempting several practice mazes, then attempt test mazes; they are scored for how many "errors" they make, which occur when the rats deviate from the correct path by crossing one of the dotted lines (see Figure 21). Now, we cannot apply the methods of agreement or difference here; genes and the environment are *permanent causes* and cannot be eliminated. We can, however, seek out or even create *variation* in these causes. For instance, we might selectively breed rats over time so as to create separate bright and dull strains of rats, with better or worse spatial memory. We might also vary the environments of the rats, creating an enriched (E), a normal (N), and a restricted environment (R) in which the rats will be reared from birth.

[60] Admittedly, this is an oversimplification of Mill's discussion. Mill also introduced the joint method of agreement and difference, the method of residues, and the problem of the intermixture of effects. However the elements introduced above will suffice to discuss the experimental nature of ANOVA.

[61] When both genotypic and environmental variation are simultaneously under investigation, these two variables make the analysis a *two-way* ANOVA. If only one variable is under investigation, then it is simply an ANOVA.

Enriched cages contain ramps, slides, bells, mirrors, marbles, polished balls, etc. The restricted cages, meanwhile, contain only a food box and a water pan. The normal environments contain standard cage accoutrements.

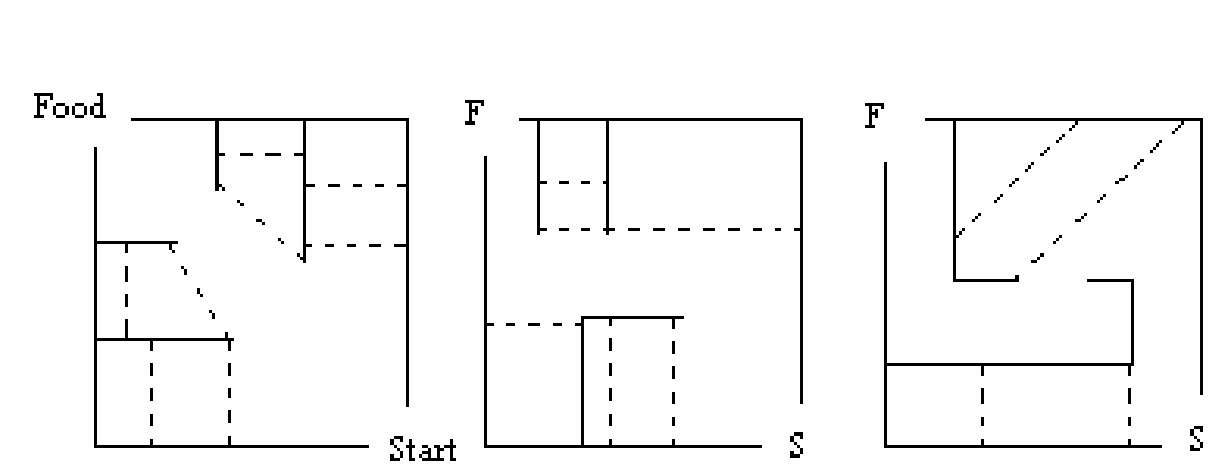

Figure 21. Sample Hebb-Williams maze test configurations.

In this hypothetical population of rats, suppose that the bright rats always on average make fewer errors than the dull rats, and rats in enriched environments always on average make fewer errors than rats in normal environments, which always on average make fewer errors than rats in restricted environments. This situation may be represented by way of a reaction norm graph (Figure 22).[62] Each strain of rat has its own reaction norm measured for mean number of errors (y-axis) at each of the various environments (x-axis).

[62] For a history of the reaction norm concept, see Sarkar (1999).

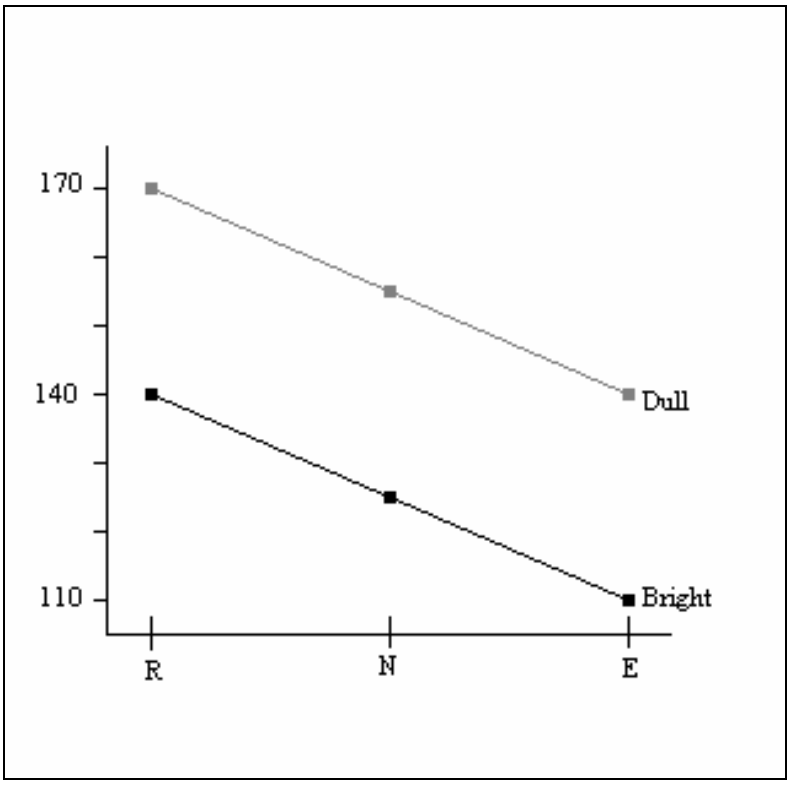

Figure 22. Hypothetical reaction norms graphing mean error scores (y-axis) for bright and dull rats in three different environments (x-axis).

Fisher's ANOVA is an application of the method of concomitant variation to a situation involving a plurality of causes in that it is designed to track the relationship between variations in genotype and variations in environment (the causes) with variation in phenotype (the effect). Fisher's innovation on Mill's method was to introduce a means of *measuring* that variation. This measuring begins by charting the error-data from Figure 22 in such a way as to calculate precisely what the difference is between each genotype for every environment and between each environment for every genotype (Table 11).

	Bright	Dull	Row Means (*R*)
Restricted	140	170	155
Normal	125	155	140
Enriched	110	140	125
Column Means (*C*)	125	155	Grand Mean (*Y*) =140

Table 11. Subgroup means for hypothetical rat population.

Now means for the genotypic and environmental measures may be calculated, providing row (*R*) and column (*C*) means along with a grand mean (*Y*) for the entire population. Utilizing the information about the population summarized in Table 11, we first calculate the total phenotypic variation (V_P) with Equation (5.1):

$$(\sum_i \sum_j (X_{ij}\text{-}Y)^2)/n \quad (5.1)$$

We measure the deviation of each subgroup mean (X_{ij}) from the grand mean (*Y*). These deviations are squared, summed, and divided by the number of subgroups (*n*) to determine the total phenotypic variation: 375. The deviation due to genotypic variation (V_G) is calculated in Equation (5.2):

$$(\sum_i (C_i - Y)^2)/c \quad (5.2)$$

We first eliminate genotypic variation by fixing on column 1 (bright) and see what the deviation from the grand mean (*Y*) is; then we eliminate genotypic variation by fixing on column 2 (dull) and see what the deviation from the grand mean is. These two deviations are squared, summed and divided by the number of columns (*c*) to determine the main effect of genotype: 225. The deviation due to environmental variation (V_E) is calculated in Equation (5.3):

$$(\sum_j (R_j - Y)^2)/r \quad (5.3)$$

We first eliminate environmental variation by fixing on row 1 (restricted) and see what the deviation from the grand mean is; then we eliminate environmental variation by fixing on row 2

(normal) and see what the deviation from the grand mean is; and then we eliminate environmental variation by fixing on row 3 (enriched) and see what the deviation from the grand mean is. These three deviations are squared, summed and divided by the number of rows (*r*) to determine the main effect of environment: 150.

This hypothetical case with its parallel reaction norms presents us with a unique quality. Focusing on Table 11, notice that as we eliminate environmental variation by fixing on each of the three environments we get the same difference between the two genotypes each time (30 errors). And as we eliminate genotypic variation by fixing on each of the two genotypes we get the same difference between the three environments each time (15 errors). There is no G×E in this hypothetical population. This is revealed visually by the fact that the reaction norms are parallel. It is revealed mathematically in the calculation we just performed. Without G×E, the total phenotypic variation (V_P) is just the sum of the "main effects": the genotypic variation (V_G) and the environmental variation (V_E):

$$V_P = V_G + V_E \quad (5.4)$$

Notice that this is precisely what we calculated in Equations 5.1, 5.2, and 5.3:

$$375 = 225 + 150$$

The situation changes quite drastically, however, when we shift our attention from the hypothetical population of rats to an *actual* population of rats. No empirical result of G×E has received more attention in the history of the nature-nurture debate than that of Cooper and Zubek (1958). Cooper and Zubek acquired both "bright" and "dull" rats; as explained above, these rats were in fact bred over time to perform better or worse in Hebb-Williams maze tests. However, all of these rats were only reared in the *normal* environments. Cooper and Zubek's interest was in what would happen if these rats were reared in either the *enriched* or the *restricted*

environments. They expected the environment to have an effect, but they also expected the bright strain to maintain its "superiority" over the dull strain across the array of environments. As Figure 23 reveals, though, that was not at all the case. In the restricted environment, the "dull" rats actually scored *fewer* errors on average than the "bright" rats (169.5 for the "dull" vs. 169.7 for the "bright"); and in the enriched environment, the "dull" rats scored only slightly more errors on average than the "bright" rats (119.7 for the "dull" vs. 111.2 for the "bright"). So it was only in the normal environment, where the "bright" rats actually earned their superior title, making the very concepts of "bright" and "dull" relative to the environments in which the rats were raised.

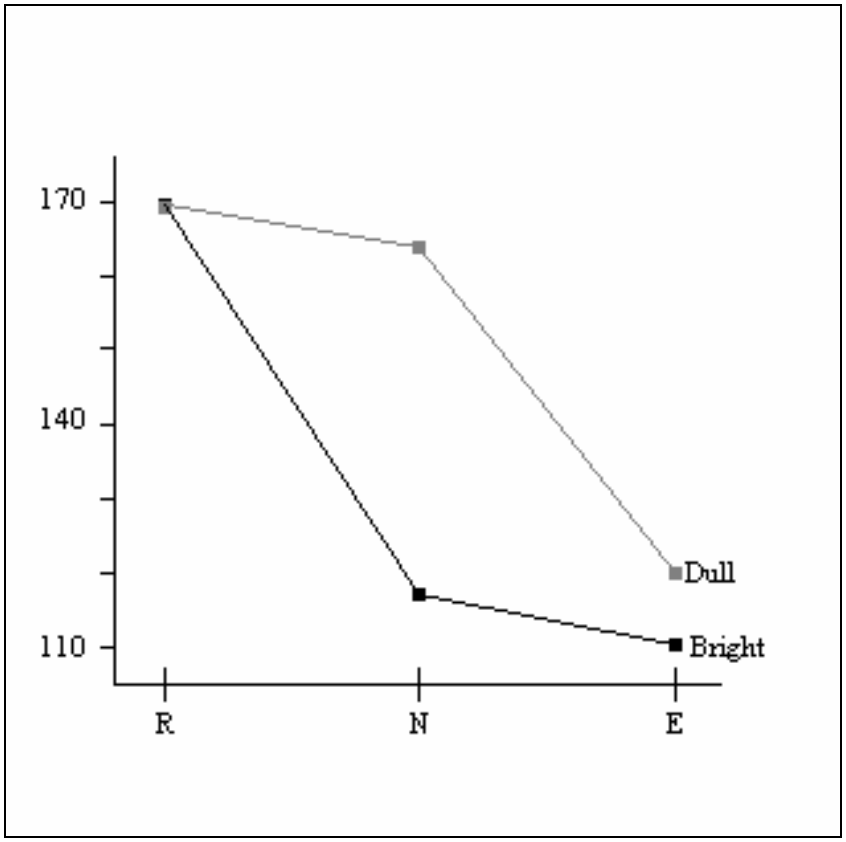

Figure 23. Actual means for Cooper and Zubek (1958) data graphed as reaction norms.

Utilizing the information about the population summarized in Table 12, the total phenotypic variation is again calculated with Equation (5.1), which equals 680; the deviation due to genotypic variation is again calculated with Equation (5.2), which equals 85; and the deviation due to environmental variation is again calculated with Equation (5.3), which equals 490. However, now we must also calculate the deviation due to G×E with Equation (5.5):

$$(\sum_i \sum_j [(X_{ij} - Y) - (C_i - Y) - (R_j - Y)]^2)/n \quad (5.5)$$

This interaction term calculates the deviation not accounted for by either the genotypic or environmental main effects, and it is 105.

	Bright	Dull	Row Means (*R*)
Restricted	169.69	169.56	169.63
Normal	117	164	140.5
Enriched	111.25	119.67	115.46
Column Means (*C*)	132.65	151.18	Grand Mean (*Y*) =141.86

Table 12. Actual subgroup means for rat population. Data from Cooper and Zubek (1958).

Notice now that when we eliminate environmental variation we actually get quite different answers. In the restricted environment the genotypes are virtually identical; in the normal environment the genotypes are separated by almost 50 errors; and in the enriched environment the genotypes nearly approach each other again. Likewise, if we fix on the dull genotype, there is little difference between the restricted and normal environments, but a huge difference between the normal and enriched environments; and if we fix on the bright genotype, there is huge difference between the restricted and normal environments, but little difference between the normal and enriched environments. The result is that total phenotypic variation no

longer is simply the sum of each source of variation if the other had been absent. Equation (5.4) will not suffice. Instead, the total phenotypic variation is the sum of the main effects of genotype and environment, *and* the variation due to interaction between genotype and environment, or Equation (5.6):

$$V_P = V_G + V_E + V_{G\times E} \quad (5.6)$$

Notice that this is precisely what we calculated in Equations 5.1, 5.2, 5.3, and 5.6:

$$680 = 85 + 490 + 105$$

The main effects do not fully account for the total phenotypic variation in spatial memory. There is a deviation from the summation formula. There is a breakdown in additivity. There is $G\times E_B$.

5.3.2. What Is the Measure Measuring?

According to the biometric interpretation of G×E, G×E is the result of the breakdown in additivity between main effects. But what is this statistical measure measuring? The biometrician might respond, "It is just measuring the breakdown in additivity." But this is only a partial answer. We must still ask the following: What *causes* this breakdown in additivity? Ironically, one of the clearest answers to this question comes from one of ANOVA's greatest critics: Richard Lewontin.

Long before Lewontin attacked Jensen's employment of ANOVA in the IQ Controversy, going so far as to claim that the statistical method was "useless" (Lewontin, 1974, 410), he actually wrote the chapter on ANOVA for the revised edition of G. G. Simpson's *Quantitative Zoology* (1960, chapter 12) (Hagen 2003). Lewontin, not yet embroiled in the heated exchange with Jensen, provided there a clear and balanced treatment of what the statistical methodology can and cannot do, along with an extensive consideration of what interaction actually *means* in

terms of the relationship between the statistical measure and the phenomenon being measured. Lewontin asked his readers to consider a population of animals sampled in different localities and at different months, making *locality* and *month* the two factors under investigation. (Focused on locality and month, the example also allows us to temporarily forget about the controversial implications that follow when the two factors are genotype and environment; while the nature-nurture debate has raged for over 100 years, the locality-month debate is far less heated…and far less distracting.) When there were two factors under investigation in a two-way ANOVA, Lewontin explained, an interaction between the two factors must be considered in addition to the main effects.

> It is the amount added to or subtracted from the basic value, arising from the particular and unique interaction of a given month with a given locality. For example, locality 5 may on the average have longer individuals than the other localities, and individuals collected in February might be larger on the average than those in other months, but it is entirely possible that individuals collected in February from locality 5 may be smaller than the average of other members of the sample. *This would presumably be due to a unique interaction of the particular locality with the particular conditions during February* (Simpson, Roe, and Lewontin 1960, 261, emphasis added).

Notice that Lewontin's last sentence is virtually identical to Hogben's third class of variability: that which "arises from the combination of a particular hereditary constitution with a particular kind of environment." (Hogben, 1933, 98)

Lewontin's example, however, does not complete the job. For we must go on to ask, what is the nature of this "particular and unique interaction"? Or, more germane, what makes the particular and unique interaction *developmental* in nature when the two factors are genotype and environment, as Hogben suggested? To see this, another example will be needed where genotype and environment actually are the factors under investigation. Fortunately, we can use an empirical example now familiar: Cooper and Zubek's (1958) study. As I pointed out above,

Cooper and Zubek's empirical results are mentioned in virtually all discussions of G×E in the confines of the nature-nurture debate. What rarely gets mentioned in these discussions, however, is Cooper and Zubek's *own* discussion of their results. After identifying the case of G×E, Cooper and Zubek did not attempt to eliminate the nuisance with a transformation of scale. Instead, they focused their attention on what causal mechanisms were responsible for the breakdown in additivity: "What physiological mechanism or mechanisms underlie these changes in learning ability?," Cooper and Zubek asked (ibid, 162). The mechanism that Cooper and Zubek considered was that proposed by Donald Hebb (1949), who argued that neural cell assemblies were built up over time (and especially during infancy) by varied stimulation coming through varied sensory pathways.[63] Applying this postulated mechanism to their own study, Cooper and Zubek offered the following explanation: In the normal environment, the level of stimulation was sufficient to permit the building of cell assemblies in the brains of the bright rats, but this level of stimulation did not meet the threshold needed to build cell assemblies in the dull rats. In the restricted environment, the level of stimulation was so low that it was inadequate for cell assembly construction in the normally bright rats, and so their error scores soared, but the dull rats were not further incapacitated since the level of stimulation provided by the normal environment was already below their threshold for the construction of the cell assemblies. Finally, in the enriched environment, the level of stimulation was far above the threshold needed by the dull rats, and so they showed a marked improvement, while the bright rats showed little

[63] Cooper and Zubek's appeal to Hebb's work was no surprise. The maze-test employed by Cooper and Zubek was designed by Hebb (Hebb and Williams 1946); Hebb actually undertook an experiment similar to Cooper and Zubek's 10 years earlier by taking several rats home from his lab to let his daughters raise them and to then see how well they subsequently performed in the maze-test compared to lab-reared rats (Hebb 1947); Hebb was an academic mentor to both Cooper and Zubek (Cooper, personal communication); and Hebb was the one person thanked by Cooper and Zubek in the acknowledgments section of their publication.

improvement because the extra stimulation was superfluous, that provided by the normal environment being adequate for the building of cell assemblies (ibid, 163).

The neurobiological accuracy of Cooper and Zubek's explanation is not particularly relevant to our discussion, although work on long-term potentiation (LTP) is beginning to bear out their account. What is relevant to our discussion is the fact that Cooper and Zubek's explanation of differences in learning ability was *developmental* in nature. The differences in genotype between the bright and dull rats did have a slight effect on total variation. The differences in environment also clearly had an effect on total variation. What differences accounted for the variation due to G×E? Cooper and Zubek attempted to answer this: A stimulating environment and the genotypically-shaped construction of cell assemblies were interdependent in such a way during individual development such that, in addition to differences in the main effects of genotype and environment, there were also differences resulting from unique combinations of genotype and environment.

The interdependent-difference-makers concept of G×E captures the causal-mechanical explanation proposed by Cooper and Zubek, as well as the relationship between $G\times E_B$ and $G\times E_D$ in this example. The genotype and environmental stimulation were both difference-makers in the development of spatial memory that took different values in the population of rats. As Cooper and Zubek discovered, these difference-makers were *not* independent of each other. They were interdependent. That is, the difference that each difference-maker made was itself dependent upon the difference made by the other difference-maker. This interdependent-difference-makers concept of G×E is just a more general, causal-mechanical interpretation of $G\times E_D$. Cooper and Zubek identified an instance of Hogben's third class of variability: In addition to genotypic variation and environmental variation, there was variation resulting from differences in unique

combinations of a particular genotype and a particular level of stimulation as they had the opportunity to interact during the process of development. Turning to the biometric contribution to the investigation, the variance attributable to G×E ($V_{G\times E}$) was 105 in the Cooper-Zubek study. This was a measure of the breakdown in additivity between the main effects of genotype and environment. But it was also a measure of the interdependence of the two difference-makers that took different values in the rat population.

Ultimately, then, it is in this sense that $G\times E_B$ and $G\times E_D$ may be integrated under the interdependent-difference-makers concept of G×E. G×E results from differences in unique, developmental combinations of genotype and environment when both variables are difference-makers in development that naturally take different values and the difference that each variable makes is itself dependent upon the difference made by the other variable; a breakdown in additivity between main effects is a measure of this interdependence of difference-makers that naturally take different values. Another way to think about the relationship is to see that the interdependent-difference-makers concept of G×E is just a more general, causal-mechanical reinterpretation of $G\times E_D$, of which $G\times E_B$ is a statistical measure.

It is important to conclude by noting that I am not alone in the conclusion that I have reached here. Douglas Wahlsten, like the critics of quantitative behavioral genetics, attacks the implications drawn from ANOVA and particularly the attempts to eradicate any interaction effects with an eye towards main effects and the heritability estimates that can be derived from them. But he does *not* follow the critics and thereby conclude that ANOVA is useless for elucidating the causal mechanisms of individual development. Instead, he suggests that ANOVA is at its best when it is detecting interaction effects (and at its worse when it is eliminating interaction effects in favor of seeking out heritability estimates) precisely because of the insights

given for understanding development when an interaction effect is found. “For those who wish to learn how development actually works,” he writes in an influential target article, “wholesale and ad hoc testing of various transformations [of scale] for the express purpose of getting rid of H×E interaction is counterproductive, because the shape of the functional relationship between variables provides a valuable clue to their causal connections.” (Wahlsten 1990, 118; see also Wahlsten 1994, 2000). Wahlsten’s arguments are largely methodological in nature, dissecting the method of ANOVA itself and identifying weakness in it, such as its inability to detect G×E without a sufficiently large sample size. In reaching the same conclusion as Wahlsten about G×E, I see my analysis above concerning the interdependence of difference-makers in development as providing a philosophical account of causation that acts as a base to support his methodological evaluations.

5.4. Conclusion

In this chapter, I brought together the various threads of the earlier chapters. In Chapters 2 and 3 I introduced the divide between the biometric and the developmental research traditions as well as the corresponding divide between the biometric ($G\times E_B$) and the developmental ($G\times E_D$) concepts of G×E. In Chapter 4 I introduced a general framework for integrating the biometric and the developmental research traditions via the concept of mechanism differences; this integrative model was contrasted with the commonly made claim that the research traditions are isolated at different levels of analysis. I did not, however, discuss $G\times E_B$ or $G\times E_D$ in that chapter. The purpose of this chapter has been to apply that general framework to the divide between $G\times E_B$ and $G\times E_D$.

I undertook this application by showing how $G\times E_B$ and $G\times E_D$ could be integrated under the interdependent-difference-makers concept of G×E. To remind ourselves, $G\times E_B$ was defined as the result of the breakdown in additivity between genotypic and environmental sources of variation, which is measured by a statistical methodology such as the analysis of variance. And $G\times E_D$ was defined as the result of differences in unique, developmental combinations of genotype and environment. In the last chapter, the interdependent-difference-makers concept of G×E was introduced; on this account, G×E was defined as the result of the interdependence of difference-makers in development that take different values in the natural world. The integrative relationship looked like this: G×E results from differences in unique, developmental combinations of genotype and environment when both variables are difference-makers in development that naturally take different values and the difference that each variable makes is itself dependent upon the difference made by the other variable, and a breakdown in additivity between main effects is a measure of this interdependence of difference-makers in development that take different values in the natural world. In the terminology of $G\times E_B$ and $G\times E_D$: $G\times E_B$ is a statistical measure of $G\times E_D$, which can itself be understood in more general, causal-mechanical terms as the result of the interdependence of difference-makers that take different values in the natural world.

6. CONCLUSION

Despite the widely endorsed "interactionist credo" (Kitcher 2001, 398), the nature-nurture debate remains a quagmire of epistemological and methodological disputes over causation, explanation, and the concepts employed therein. I hope I have conveyed the nature of this quagmire in the previous chapters…as well as a potential way out of the mess. In each chapter, I tried to make explicit the theses about the quagmire as well as the theses about the way out. The quagmire was introduced in Chapters 2 and 3 (different research traditions, different concepts of G×E), and the way out was furnished in Chapters 4 and 5 (difference mechanisms, interdependent-difference-makers concept of G×E). There have, however, been several *implicit* theses of the dissertation as well, and I would like to make those explicit now. This will also give me a chance to point to some new directions to which the research lends itself.

6.1. Causes of Controversy

What causes controversy? For the nature-nurture debate, the answer to this question has largely been supplied by appeals to socio-political motivations, both from the disputants themselves and from historians and philosophers reflecting on the debate. Fisher was accused of bias against particular classes; Hogben was accused of ideological socialism; Jensen was accused of racism; Lewontin was accused of dogmatic Marxism. To a certain extent, these socio-political motivations are indisputable and uncontroversial. Fisher was an avowed eugenicist; Lewontin openly embraced Marxism. Indeed, it would be unusual if these political motivations did *not* shape how disputants responded to the issues, since the nature-nurture debate has clear implications for how society conceives of and responds to differences among individuals. The problem arises, however, when admitting to the existence of socio-political motivations leads to

assuming that *all* there is to the controversy is the biases of disputants. Biologists in the developmental tradition have simply taken it as obvious that G×E is a developmental phenomenon. When Jensen did not see this blatant truth, his opponents concluded that he must be motivated by irrational biases, such as racism. Biologists in the biometric tradition have simply taken it as obvious that G×E is a statistical phenomenon. When Lewontin did not see this blatant truth, his opponents concluded that he must be motivated by irrational biases, such as dogmatic Marxism.

I have tried to draw attention to a different cause of the controversy. There are substantive epistemological, methodological, and conceptual issues involved in the debates over G×E. However, these issues are far from obvious. When disagreement arises, it is much easier to accuse your opponent of blinded racism then to take a step back and reflect on differences in problem, differences in methodology, differences in questions, differences in concepts. As I showed, these components are embedded in particular research traditions, and it is no easy task to step out of a particular research tradition and consider the virtues and vices of another. But if research traditions are to integrate, as a growing number of scientists and philosophers of science are suggesting, then it is precisely this attention to the epistemological, methodological, and conceptual issues that must be afforded. A relationship between the traditions cannot be developed until the differences between the various components of the traditions are disentangled, and these differences cannot be disentangled when it is assumed that socio-political motivations are the sole cause of the controversy.

One of the ironies of the dissertation is that the disputant perhaps most motivated by socio-political concerns and most likely to attribute socio-political biases to his opponent was also the disputant with the most to offer in the way of considering the epistemological,

methodological, and conceptual issues. Hogben, having risen above the British class system of his day, was a proud socialist, and he was eager to infuse the lessons of his politics into his scientific research. His class-ascendancy also made him incredibly paranoid, and he held grudges for a lifetime. But in spite of all that, Hogben recognized that his disagreement with Fisher was largely about causation, explanation, and the concepts employed therein. "When you speak of the contribution of hereditable and non-hereditable causes of variance in a population," he asked Fisher, "what exactly do you mean?" How was Hogben able to recognize this epistemological divide, where subsequent disputants have not? I doubt there is a single, simple answer to the question. But it is certainly worth noting that Hogben, unlike any other individual mentioned in the earlier chapters, was trained in and then undertook research in both the biometric and developmental research traditions. This experience in both research traditions certainly put Hogben in the best position to compare and contrast the epistemological underpinnings of each.

6.2. Making a Difference with Difference Mechanisms

So what now? I dealt with one aspect of the nature-nurture debate in the dissertation—the debates over G×E. There is reason to believe, however, that the difference mechanisms solution has broader application, both within behavioral genetics and beyond. Within behavioral genetics, there is still room for a general framework of causal-mechanical explanation within the discipline. The biometric methodologies have advanced well-beyond ANOVA, and so there is still work to be done showing what role, say, structural equation modeling plays in the elucidation of difference mechanisms. Or, the solution could be extended to capture genotype-environment correlation…perhaps an account of correlated difference-makers in development

that take different values in the natural world. One future direction of the research, then, will involve extending the difference mechanisms solution to behavioral genetics more broadly.

As I suggested in Chapter 4, this extension need not stop with behavioral genetics. The common denominator of difference mechanisms is applicable wherever both regularity and variation are under investigation. That is, difference mechanisms are applicable wherever both proximate how-questions and ultimate why-questions are being asked. Difference mechanisms may offer a general relationship between proximate and ultimate biology, and it will be worth testing this hypothesis by extending it to other disciplines that bridge the divide, such as evolutionary developmental biology.

BIBLIOGRAPHY

Aby, Stephen H., and Martha J. McNamara (Eds.). 1990. *The IQ Debate: A Selective Guide to the Literature*. New York: Greenwood Press.

Allen, Garland E. 1986. "The Eugenics Record Office, Cold Spring Harbor, 1910-1940: An Essay in Institutional History", *Osiris*:225-264.

Anderson, T.W. 1996. "R.A. Fisher and Multivariate Analysis", *Statistical Science* 11:20-34.

Ariew, Andre. 2003. "Ernst Mayr's 'Ultimate/Proximate' Distinction Reconsidered and Reconstructed", *Biology and Philosophy* 18: 553-565.

Aronson, J. K. 1985. *An Account of the Foxglove and Its Medical Uses, 1785-1985*. Oxford: Oxford University Press.

Barkan, Elazar. 1991. "Reevaluating Progressive Eugenics: Herbert Spencer Jennings and the 1924 Immigration Legislation", *Journal of the History of Biology* 24:91-112.

Beatty, John. 1994. "The Proximate/Ultimate Distinction in the Multiple Careers of Ernst Mayr", *Biology and Philosophy* 9: 333-356.

Beatty, John. 1995. "The Evolutionary Contingency Thesis", in G. Wolters and J. G. Lennox (Eds.), *Concepts, Theories, and Rationality in the Biological Sciences*. Pittsburgh: University of Pittsburgh Press, pp. 45-81.

Beatty, John. 1997. "Why Do Biologists Argue Like They Do?", *Philosophy of Science* 64: S432-S443.

Bechtel, William and Robert C. Richardson. 1993. *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*. Princeton: Princeton University Press.

Bechtel, William and Adele Abrahamsen. 2005. "Explanation: A Mechanist Alternative", *Studies in the History and Philosophy of Biology and the Biomedical Sciences* 36: 421-441.

Bengel, Dietmar, Dennis L. Murphy, Anne M. Andrews, et al. 1998. "Altered Brain Serotonin Homeostasis and Locomotor Insensitivity to 3,4-Methylenedioxymethamphetamine ("Ecstacy") in Serotonin Transporter-Deficient Mice", *Molecular Pharmacology* 53: 649-655.

Bennett, J.H. (Ed.). 1983. *Natural Selection, Heredity, and Eugenics: Including Selected Correspondence of R.A. Fisher with Leonard Darwin and Others*. Oxford: Clarendon Press.

Bennett, J.H. (Ed.). 1990. *R. A. Fisher, Statistical Inference and Analysis: Selected Correspondence of R. A. Fisher*. Oxford: Oxford University Press.

Bergmann, Uri. 2000. "In Your Mind's Eye: Speculations on the Neurobiology of Eye Movement Desensitization and Reprocessing (EMDR)", *New Therapist* 9.

Blacker, C. P. 1952. *Eugenics: Galton and After*. Cambridge, Mass: Harvard University Press.

Block, Ned. 1995. "How Heritability Misleads About Race", *Cognition* 56: 99-128.

Block, Ned and Gerald Dworkin. 1976. "IQ, Heritability and Inequality", in Ned Block and Gerald Dworkin (eds.), *The IQ Controversy*. New York: Pantheon, pp. 410-540.

Bouchard, Thomas J., and Nancy L. Segal. 1985. "Environment and IQ", in B. B. Wolman (Ed.), *Handbook of Intelligence: Theories, Measurements, and Applications*. New York: John Wiley and Sons, pp. 391-464.

Box, Joan Fisher. 1978. *R.A. Fisher: The Life of a Scientist*. New York: John Wiley and Sons.

Brandon, Robert 1997. "Does Biology Have Laws? The Experimental Evidence", *Philosophy of Science* 64: S444-S457.

Burt, Cyril. 1955. "The Evidence for the Concept of Intelligence", *British Journal of Educational Psychology* 25: 158-177.

Burt, Cyril. 1958. "The Inheritance of Mental Ability", *American Psychologist* 13: 1-15.

Burt, Cyril. 1966. “The Genetic Determination of Differences in Intelligence: A Study of Monozygotic Twins Reared Together and Apart”, *British Journal of Psychology* 57: 137-153.

Caspi, Avshalom, Joseph McClay, Terrie E. Moffitt, et al. 2002. “Role of Genotype in the Cycle of Violence in Maltreated Children”, *Science* 297: 851-854.

Caspi, Avshalom, Karen Sugden, Terrie E. Moffitt, et al. 2003. “Influence of Life Stress on Depression: Moderation by a Polymorphism in the 5-HTT Gene”, *Science* 301: 386-389.

Caspi, Avshalom, Terrie E. Moffitt, Mary Cannon, et al. 2005. “Moderation of the Effect of Adolescent-Onset Cannabis Use in Adult Psychosis by a Functional Polymorphism in the Catechol-O-Methyltransferase Gene: Longitudinal Evidence of a Gene X Environment Interaction”, *Biological Psychiatry* 57: 1117-1127.

Caspi, Avshalom and Terrie E. Moffitt. 2006. “Gene-Environment Interactions in Psychiatry: Joining Forces with Neuroscience”, *Nature Reviews Neuroscience* 7: 583-590.

Chase, Allan. 1980. *The Legacy of Malthus: The Social Costs of the New Scientific Racism*. Urbana: University of Illinois Press.

Cochran, William G. 1980. “Fisher and the Analysis of Variance”, in S. E. Fienberg and D. V. Hinkley (Eds.), *R.A. Fisher: An Appreciation*. New York: Springer-Verlag, pp. 17-34.

Cooper, Roderick M. and John P. Zubek. 1958. “Effects of Enriched and Restricted Early Environments on the Learning Ability of Bright and Dull Rats”, *Canadian Journal of Psychology* 12(3): 159-164.

Craver, Carl F. 2001. “Structures of Scientific Theories”, in P. K. Machamer and M. Silberstein (Eds.), *Blackwell Guide to the Philosophy of Science*. Oxford: Blackwell.

Craver, Carl F. 2002. “Interlevel Experiments and Multilevel Mechanisms in the Neuroscience of Memory”, *Philosophy of Science* 69: S83-S97.

Craver, Carl F. 2005. “Beyond Reduction: Mechanisms, Multifield Integration and the Unity of Neuroscience”, *Studies in the History and Philosophy of Biology and the Biomedical Sciences* 36: 373-395.

Craver, Carl F. and Lindley Darden. 2001. “Discovering Mechanisms in Neurobiology: The Case of Spatial Memory”, in P. K. Machamer, R. Grush, and P. McLaughlin (Eds.), *Theory and Method in the Neurosciences*. Pittsburgh: Pittsburgh University Press, pp. 112-137.

Crews, F. A. E., W. J. Dakin, J. Heslop Harrison, Lancelot T. Hogben, Julian S. Huxley, J. Johnston, F. H. A. Marshall, Guy C. Robson, A. M. Carr Saunders, and J. MacLean Thompson. 1923. “The British Journal of Experimental Biology”, *Science* 58: 102.

Cronbach, Lee J. and Paul E. Meehl. 1955. “Construct Validity in Psychological Tests”, *Psychological Bulletin* 52: 281-302.

Daniels, Norman. 1974. “IQ, Heritability, and Human Nature”, *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 143-180.

Darden, Lindley, and Nancy Maull. 1977. “Interfield Theories”, *Philosophy of Science* 44: 43-64.

Dijksterhuis, E. J. 1961. *The Mechanization of the World Picture*. Translated by C. Dikshoorn. Oxford: Clarendon Press.

Dobzhansky, Theodosius. 1973. *Genetic Diversity and Human Equality*. New York: Basic Books, Inc.

Dobzhansky, Theodosius, and Boris Spassky. 1944. “Genetics of Natural Populations. XXXIV. Adaptive Norm, Genetic Load, and Genetic Elite in *Drosophila pseudoobscura*”, *Genetics* 48: 1467-1485.

Dowe, Phil. 2000. *Physical Causation*. Cambridge: Cambridge University Press.

Downes, Stephen. 2004. "Heredity and Heritability", in Edward N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. URL = http://plato.stanford.edu/archives/fall2004/entries/heredity/.

Eckberg, Douglas Lee. 1979. *Intelligence and Race: The Origins and Dimensions of the IQ Controversy*. New York: Praeger.

Economist. 2002. "The Origins of Violence: Nurturing Nature", 364: 71.

Edwards, A. W. F. 1990. "Fisher, W, and the Fundamental Theorem", *Theoretical Population Biology* 38: 276-284.

Eysenck, Hans J. 1971. *The IQ Argument: Race, Intelligence and Education*. LaSalle, Ill: Open Court Publishing Co.

Eysenck, Hans J. 1973. *The Inequality of Man*. London: Temple Smith.

Eysenck, Hans J., and David W. Fulker. 1979. *The Structure and Measurement of Intelligence*. Berlin: Springer-Verlag.

Erlingsson, Steindor. 2005. *The Rise of Experimental Zoology in Britain in the 1920's: Hogben, Huxley, Crew, and the Society for Experimental Biology*. Unpublished Dissertation. Manchester: University of Manchester.

Falconer, D. S. and T. F. C. Mackay. 1996. *Introduction to Quantitative Genetics*. Essex, Longman Group Limited.

Feldman, Marcus W., and Richard C. Lewontin. 1975. "The Heritability Hang-up", *Science* 190: 1163-1168.

Finney, D. J. 1964. "Sir Ronald Fisher's Contributions to Biometric Statistics", *Biometrics* 20:322-329.

Fisher, Ronald A. 1915. “Frequency Distribution of the Values of the Correlation Coefficient in Samples from an Indefinitely Large Population”, *Biometrika* 10:507-521.

Fisher, Ronald A. 1918. “The Correlation between Relatives on the Supposition of Mendelian Inheritance”, *Transactions of the Royal Society of Edinburgh* 52:399-433.

Fisher, Ronald A. 1924. “The Biometrical Study of Heredity”, *Eugenics Review* 16:189-210.

Fisher, Ronald A. 1925. *Statistical Methods for Research Workers*. Edinburgh: Oliver and Boyd Ltd.

Fisher, Ronald A. 1932. “Review of Lancelot Hogben’s *Genetic Principles in Medicine and Social Science*”, *Health and Empire* 7:147-150.

Fisher, Ronald A. 1958. *The Genetical Theory of Natural Selection*. Dover: New York.

Fisher, Ronald A., and W. A. Mackenzie. 1923. “Studies in Crop Variation. II. The Manurial Response of Different Potato Varieties”, *Journal of Agricultural Science* 13:311-320.

Gillispie, Charles Coulston. 1960. *The Edge of Objectivity*. Princeton: Princeton University Press.

Glennan, Stuart. 2002. “Rethinking Mechanistic Explanation”, *Philosophy of Science* 69: S342-S353.

Gottlieb, Gilbert. 2003. “On Making Behavioral Genetics Truly Developmental”, *Human Development* 46: 337-355.

Griffiths, Paul E. and James Tabery. Forthcoming. “Behavioral Genetics and Development: Historical and Conceptual Causes of Controversy”, *New Ideas in Psychology*.

Gurdon, John B., and Nick Hopwood. 2000. “The Introduction of *Xenopus laevis* into Developmental Biology: Of Empire, Pregnancy Testing and Ribosomal Genes”, *International Journal of Developmental Biology* 44:43-50.

Hagen, Joel. 2003. "The Statistical Frame of Mind in Systematic Biology from *Quantitative Zoology* to *Biometry*", *Journal of the History of Biology* 36: 353-384.

Haldane, J. B. S. 1932. "A Programme for Human Genetics: A Review of Lancelot Hogben's *Genetic Principles in Medicine and Social Science*", *Nature* 129: 345-346.

Hall, Marie. 1952. "The Establishment of the Mechanical Philosophy", *Osiris* 10: 412-541.

Hamer, Dean. 2004. "Rethinking Behavior Genetics", *Science* 298: 71-72.

Hamilton, Andrew. 2006. "Review of Ernst Mayr's What Makes Biology Unique?: Considerations on the Autonomy of a Scientific Discipline", *Philosophy of Science* 73: 255-257.

Hariri, Ahmad R., Venkata S. Mattay, Alessandro Tessitore, et al. 2002. "Serotonin Transporter Genetic Variation and the Response of the Human Amygdala", *Science* 297: 400-403.

Hariri, Ahmad R. and Andrew Holmes. 2006. "Genetics of Emotional Regulation: The Role of the Serotonin Transporter in Neural Function", *TRENDS in Cognitive Science* 10: 182-191.

Hebb, Donald O. 1947. "The Effects of Early Experience on Problem Solving at Maturity", *American Psychologist* 2: 306-307.

Hebb, Donald O. 1949. *The Organization of Behavior*. New York: Wiley.

Hebb, Donald O. and K. Williams. 1946. "A Method of Rating Animal Intelligence", *Journal of General Psychology* 34: 59-65.

Hempel, Carl and P. Oppenheim .1948. "Studies in the Logic of Explanation", *Philosophy of Science* 15: 135-175.

Hernandez, Lyla M. and Dan G. Blazer (Eds.). 2006. *Genes, Behavior, and the Social Environment: Moving Beyond the Nature/Nurture Debate*. Washington, D.C.: The National Academic Press.

Herrnstein, Richard J. 1971. "I.Q.", *The Atlantic Monthly* 228: 43-64.

Herrnstein, Richard J. 1973. *I.Q. in the Meritocracy*. Boston: Little, Brown and Company.

Hogben, Lancelot. 1927. *Principles of Evolutionary Biology*. Cape Town: Juta.

Hogben, Lancelot. 1931a. *The Nature of Living Matter*. New York: Alfred A. Knopf.

Hogben, Lancelot. 1931b. "The Foundations of Social Biology", *Economica* 31:4-24.

Hogben, Lancelot. 1931c. "The Genetic Analysis of Familial Traits. I. Single Gene Substitutions", *Journal of Genetics* 25:97-112.

Hogben, Lancelot. 1932a. *Genetic Principles in Medicine and Social Science*. New York: Alfred A. Knopf.

Hogben, Lancelot. 1932b. "The Genetic Analysis of Familial Traits. II. Double Gene Substitutions, with Special Reference to Hereditary Dwarfism", *Journal of Genetics* 25:211-240.

Hogben, Lancelot. 1932c. "The Genetic Analysis of Familial Traits. III. Matings Involving One Parent Exhibiting a Trait Determined by a Single Recessive Gene Substitution with Special Reference to Sex-linked Conditions", *Journal of Genetics* 25:293-314.

Hogben, Lancelot. 1933a. *Nature and Nurture, Being the William Withering Memorial Lectures*. London: George Allen and Unwin Ltd.

Hogben, Lancelot. 1933b. "The Limits of Applicability of Correlation Technique in Human Genetics", *Journal of Genetics* 27:379-406.

Hogben, Lancelot. 1937. *Mathematics for the Million*. New York: W. W. Norton and Company, Inc.

Hogben, Lancelot. 1938. *Science for the Citizen: A Self-Educator Based on the Social Background of Scientific Discovery*. New York: Alfred A. Knopf.

Hogben, Lancelot. 1998. *Lancelot Hogben, Scientific Humanist: An Unauthorized Autobiography*. Suffolk: Merlin Press.

Hogben, Lancelot T. and Frank R. Winton. 1922a. "The Pigmentary Effector System. I.—Reaction of Frog's Melanophores to Pituitary Extracts", *Proceedings of the Royal Society of London, Series B* 93: 318-329.

Hogben, Lancelot T. and Frank R. Winton. 1922b. "The Pigmentary Effector System. II", *Proceedings of the Royal Society of London, Series B* 94: 151-162.

Hogben, Lancelot T. and Frank R. Winton. 1923. "The Pigmentary Effector System. III.—Colour Response in the Hypophysectomised Frog." *Proceedings of the Royal Society of London, Series B* 95: 15-31.

Hume, David. 1993 [1777]. *An Enquiry Concerning Human Understanding*. Indianapolis: Hackett Publishing Company.

Huxley, Julian. 1916. "Bird-watching and Biological Science", *Auk* 33: 142-161, 256-270.

Huxley, Julian. 1932. "Eugenics: A Review of Lancelot Hogben's *Genetic Principles in Medicine and Social Science*", *Eugenics Review* 23: 341-344.

Huxley, Julian S. and Lancelot T. Hogben. 1922. "Experiments on Amphibian Metamorphosis and Pigment Responses in Relation to Internal Secretions", *Proceedings of the Royal Society of London, Series B* 93: 36-53.

Jensen, Arthur R. 1969. "How Much Can We Boost IQ and Scholastic Achievement?", *Harvard Educational Review* 39: 1-123.

Jensen, Arthur R. 1970. “Race and the Genetics of Intelligence: A Reply to Lewontin”, *Bulletin of the Atomic Scientist* 26: 17-23.

Jensen, Arthur R. 1972. “The IQ Controversy: A Reply to Layzer”, *Cognition* 1: 427-452.

Jensen, Arthur R. 1973. *Educability and Group Differences*. New York: Harper and Row, Publishers.

Jensen, Arthur R. 1975. “The Meaning of Heritability in the Behavioral Sciences”, *Educational Psychologist* 11: 171-183.

Jensen, Arthur R. 1976. “Heritability of IQ”, *Science* 194: 6-8.

Jensen, Arthur R. 1981. *Straight Talk about Mental Tests*. London: Methuen.

Jinks, J. L., and David W. Fulker. 1970. “Comparison of the Biometrical Genetical, MAVA, and Classical Approaches to the Analysis of Human Behavior”, *Psychological Bulletin* 73: 311-349.

Johnstone, D.J. 1987. “Tests of Significance Following R.A. Fisher”, *British Journal for the Philosophy of Science* 38:481-499.

Kaplan, Jonathan Michael. 2000. *The Limits and Lies of Human Genetic Research*. New York: Routledge.

Kearsey, M. J. and H. S. Pooni. 1996. *The Genetical Analysis of Quantitative Traits*. London, Chapman and Hall.

Kendler, Kenneth S. 2005. “Toward a Philosophical Structure for Psychiatry”, *American Journal of Psychiatry* 162: 433-440.

Kevles, Daniel J. 1995. *In the Name of Eugenics: Genetics and the Uses of Human Heredity*. Cambridge, Mass: Harvard University Press.

Kitcher, Philip. 2001. “Battling the Undead: How (and How Not) to Resist Genetic Determinism”, in R. S. K. Singh, B. Costas, Diane B. Paul, and John Beatty (Eds.), *Thinking About Evolution: Historical, Philosophical, and Political Perspectives*. Cambridge: Cambridge University Press. pp. 396-414.

Krafka, Joseph. 1920. “The Effect of Temperature Upon Facet Number in the Bar-eyed Mutant of Drosophila”, *Journal of General Physiology* 2:409-464.

Kuhn, Thomas. 1962. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Lakatos, Imre. 1970. “Falsification and the Methodology of Scientific Research Programmes”, in Imre Lakatos and Alan Musgrave (Eds.), *Criticism and the Growth of Knowledge*. Cambridge: Cambridge University Press, pp. 91-196.

Laudan, Larry. 1977. *Progress and Its Problems*. Berkeley: University of California Press.

Layzer, David. 1972a. “Science or Superstition? (A Physical Scientist Looks at the IQ Controversy)”, *Cognition* 1: 265-299.

Layzer, David. 1972b. “Jensen’s Reply: The Sounds of Silence”l *Cognition* 1: 453-473.

Layzer, David. 1974. “Heritability Analyses of IQ Scores: Science or Numerology”l *Science* 183: 1259-1266.

Leonardo, E. D. and Rene Hen. 2006. “Genetics of Affective and Anxiety Disorders”, *Annual Review of Psychology* 57: 117-137.

Lesch, Klaus-Peter, Dietmar Bengel, Armin Heils, et al. 1996. “Association of Anxiety-Related Traits with a Polymorphism in the Serotonin Transporter Gene Regulatory Region”, *Science* 274: 1527-1531.

Levin, Michael. 1997. *Why Race Matters: Race Differences and What They Mean*. Westport, Connecticut: Praeger.

Lewis, David. 1973. “Causation”, *The Journal of Philosophy* 70(17): 556-567.

Lewontin, Richard C. 1970a. “Race and Intelligence”, *Bulletin of the Atomic Scientist* 26: 2-8.

Lewontin, Richard C. 1970b. “Further Remarks on Race and the Genetics of Intelligence”, *Bulletin of the Atomic Scientist* 26: 23-25.

Lewontin, Richard C. 1974. “The Analysis of Variance and the Analysis of Causes”, *American Journal of Human Genetics* 26: 400-411.

Lewontin, Richard C. 1975. “Genetic Aspects of Intelligence”, *Annual Review of Genetics* 9: 387-405.

Ludmerer, Kenneth M. 1972. *Genetics and American Society: A Historical Appraisal*. Baltimore: The Johns Hopkins University Press.

Lush, Jay L. 1937. *Animal Breeding Plans*. Ames: Collegiate Press, Inc.

Lush, Jay L. 1943. *Animal Breeding Plans*. Ames: Collegiate Press.

Lynch, Michael and Bruce Walsh. 1997. *Genetics and Analysis of Quantitative Traits*. Sunderland, Massachusetts, Sinauer Associates, Inc.

Machamer, Peter K., Lindley Darden, and Carl F. Craver. 2000. “Thinking About Mechanisms”, *Philosophy of Science* 67: 1-25.

MacKenzie, Donald A. 1981. *Statistics in Britain: 1865-1930, The Social Construction of Scientific Knowledge*. Edinburgh: Edinburgh University Press.

Mahalanobis, P. C. 1964. “Professor Ronald Aylmer Fisher”, *Biometrics* 20:238-250.

Mather, K. 1964. “R.A. Fisher’s Work in Genetics”, *Biometrics* 20:330-342.

Matthews, Michael R. 1980. *The Marxist Theory of Schooling: A Study of Epistemology and Education*. Sussex: Harvester Press.

Mayr, Ernst. 1961. “Cause and Effect in Biology”, *Science* 134: 1501-1506.

Mayr, Ernst. 1982. *The Growth of Biological Thought: Diversity, Evolution, and Inheritance*. Cambridge, Mass: Belknap Press.

Mazumdar, Pauline M. H. 1992. *Eugenics, Human Genetics, and Human Failings: The Eugenics Society, Its Sources and Its Critics in Britain*. London: Routledge.

Meaney, Michael. 2001. "Nature, Nurture, and the Disunity of Knowledge", *Annals of the New York Academy of Sciences* 935: 50-61.

Mill, John Stuart. 1974 [1843]. *A System of Logic, Ratiocinative, and Inductive: Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation*. Toronto: University of Toronto Press.

Mitchell, Sandra D. 1992. "On Pluralism and Competition in Evolutionary Explanations", *American Zoologist* 32: 135-144.

Mitchell, Sandra D. 2002. "Integrative Pluralism", *Biology and Philosophy* 17: 55-70.

Mitchell, Sandra D. 1997. "Pragmatic Laws", *Philosophy of Science* 64: S468-S479.

Mitchell, Sandra D. 2003. *Biological Complexity and Integrative Pluralism*. Cambridge: Cambridge University Press.

Moffitt, Terrie E., Avshalom Caspi, and Michael Rutter. 2005. "Strategy for Investigating Interactions Between Measured Genes and Measured Environments", *Archives of General Psychiatry* 62: 473-481.

Moran, P. A. P., and C. A. B. Smith. 1966. "Commentary on R. A. Fisher's Paper on The Correlation Between Relatives on the Supposition of Mendelian Inheritance", *Eugenics Laboratory Memoirs* 41:1-62.

Norton, B. J. 1975. "Biology and Philosophy: The Methodological Foundations of Biometry", *Journal of the History of Biology* 8:85-93.

Norton, B. J. 1978. “Fisher and the Neo-Darwinian Synthesis.” in E. G. Forbes (Ed.), *Human Implications of Scientific Advance. Proceedings of the XVth International Congress of the History of Science*. Edinburgh: Edinburgh University Press, pp. 481-494.

Norton, B. J., and E. S. Pearson. 1976. “A Note on the Background to, and Refereeing of, R. A. Fisher’s 1918 Paper ‘On the Correlation of Relatives on the Supposition of Mendelian Inheritance’”, *Notes and Records of the Royal Society of London* 31:151-162.

Parens, Erik. 2004. “Genetic Differences and Human Identities: On Why Talking about Behavioral Genetics Is Important and Difficult”, *Hastings Center Report* Supplement 34: S1-S36.

Paul, Diane B. 1995. *Controling Human Heredity: 1865 to the Present*. Atlantic Highlands, NJ: Humanities Press.

Paul, Diane B. 1998. *The Politics of Heredity: Essays on Eugenics, Biomedicine, and the Nature-Nurture Debate*. Albany: State University of New York Press.

Pearson, E. S. 1968. “Studies in the History of Probability and Statistics. XX: Some Early Correspondence between W. S. Gosset, R. A. Fisher and Karl Pearson, with Notes and Comments”, *Biometrika* 55:445-457.

Pearson, Karl, and Alice Lee. 1903. “On the Laws of Inheritance in Man: I. Inheritance of Physical Chracters”, *Biometrika* 2:357-462.

Pezawas, Lukas, Andreas Meyer-Lindenberg, Emily M. Drabant, et al. 2005. “5-HTTLPR Polymorphism Impacts Human Cingulate-Amygdala Interactions: A Genetic Susceptibility Mechanism for Depression”, *Nature Neuroscience* 8: 828-834.

Piaget, Jean. 1952. *The Origins of Intelligence in Children*. New York: International Universities Press.

Plomin, Robert, and John C. DeFries. 1976. "Heritability of IQ", *Science* 194: 10-12.

Plomin, Robert, John C. DeFries, and John C. Loehlin. 1977. "Genotype-environment Interaction and Correlation in the Analysis of Human Behavior", *Psychological Bulletin* 84: 309-322.

Plomin, Robert, John C. DeFries, and David W. Fulker. 1988. *Nature and Nurture During Infancy and Early Childhood*. Cambridge: Cambridge University Press.

Plomin, R., and S. Hershberger. 1991. "Genotype-Environment Interaction", in T. D. Wachs and Robert Plomin (Eds.), *Conceptualization and Measurement of Organism-Environment Interaction*. Washington, DC: American Psychological Association, pp. 29-43.

Plomin, Robert, John C. DeFries, Gerald E. McClearn, and Michael Rutter. 1997. *Behavioral Genetics*. New York: W.H. Freeman and Company.

Plutynski, Anya. 2006. "What Was Fisher's Fundamental Theorem of Natural Selection and What Was It For?", *Studies in History and Philosophy of Biological and Biomedical Sciences* 37: 59-82.

Porter, Theodore 2004. *Karl Pearson*. Princeton: Princeton University Press.

Preece, D.A. 1990. "R.A. Fisher and Experimental Design: A Review", *Biometrics* 46:925-935.

Provine, William B. 2001. *The Origins of Theoretical Population Genetics*. Chicago: The University of Chicago Press.

Rao, C. Radhakrishna. 1964. "Sir Ronald Aylmer Fisher-The Architect of Multivariate Analysis", *Biometrics* 20:286-300.

Rendel, J. M. 1959. "Canalization of the Scute Phenotype of Drosophila", *Evolution* 13: 425-439.

Rowe, David C. 2001. *Biology and Crime*. Los Angeles, Roxbury.

Rutter, Michael. 2006. *Genes and Behavior: Nature-Nurture Interplay Explained*. Malden, MA: Blackwell Publishing.

Salmon, Wesley. 1984. *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press.

Salmon, Wesley. 1998. *Causality and Explanation*. Oxford: Oxford University Press.

Sarkar, Sahotra. 1996. "Lancelot Hogben, 1895-1975", *Genetics* 142:655-660.

Sarkar, Sahotra. 1998. *Genetics and Reductionism*. Cambridge: Cambridge Universitiy Press.

Sarkar, Sahotra. 1999. "From the Reaktionsnorm to the Adaptive Norm: The Norm of Reaction, 1909-1960", *Biology and Philosophy* 14: 235-252.

Scarr, Sandra. 1995. "Commentary on Gottlieb's 'Some Conceptual Deficiencies in 'Developmental' Behavior Genetics'", *Human* Development 38: 154-158.

Schaffner, Kenneth F. 1993. *Discovery and Explanation in Biology and Medicine*. Chicago: The University of Chicago Press.

Sesardic, Neven. 1993. "Heritability and Causality", *Philosophy of Science* 60: 396-418.

Sesardic, Neven. 2005. *Making Sense of Heritability*. Cambridge: Cambridge University Press.

Shapere, Dudley. 1977. "Scientific Theories and Their Domains", in Frederick Suppe (Ed.), *The Structure of Scientific Theories*. Urbana: University of Illinois Press, pp. 518-565.

Sherman, Paul W. 1988. "The Levels of Analysis", *Animal Behavior* 36: 616-619.

Sherman, Paul W. 1989. "The Clitoris Debates and the Levels of Analysis", *Animal Behavior* 37: 697-698.

Shih, J. C., K. Chen, et al. 1999. "MONOAMINE OXIDASE: From Genes to Behavior", *Annural Review of Neuroscience* 22: 197-217.

Shockley, William. 1972. "The Apple-of-God's-Eye Obsession", *Humanist* 32: 16-17.

Shockley, William. 1978. “Humanitarianism Gone Berserk?”, in *Dallas Times Herald*, September 17, Dallas, Texas.

Simpson, George Gaylord, Anne Roe, Richard C. Lewontin. 1960. *Quantitative Zoology, Revised Edition*. New York: Harcourt, Brace and Company.

Sinnott, Edmund W., Leslie C. Dunn, and Theodosius Dobzhansky. 1958. *Principles of Genetics*. New York: McGraw-Hill Book Company.

Skipper, Robert A. Jr. 2002. “The Persistence of the R. A. Fisher-Sewall Wright Controversy”, *Biology and Philosophy* 17:341-367.

Sober, Elliott. 1984. *The Nature of Selection*. Cambridge, Mass.: MIT Press.

Sober, Elliott. 1997. “Two Outbreaks of Lawlessness in Recent Philosophy of Biology”, *Philosophy of Science* 64: S458-S467.

Soloway, Richard A. 1990. *Demography and Degeneration: Eugenics and the Declining Birthrate in Twentieth-Century Britain*. Chapel Hill: The University of North Carolina Press.

Spearman, Charles. 1904. “‘General Intelligence,’ Objectively Determined and Measured”, *American Journal of Psychology* 15: 201-293.

Speed, T. P. 1987. “What is an Analysis of Variance?”, *The Annals of Statistics* 15: 885-941.

Sokal, Robert R. and F. James Rohlf. 1995. *Biometry: The Principles and Practice of Statistics in Biological Research*. New York: W. H. Freeman.

Surbey, M. K. 1994. “Discussion: Why Expect a Horse to Fly?, reply to Wahlsten”, *Canadian Psychology* 35: 261-264.

Tabery, James G. 2004. “Synthesizing Activities and Interactions in the Concept of a Mechanism”, *Philosophy of Science* 71: 1-15.

Tabery, James G. 2004. "The 'Evolutionary Synthesis' of George Udny Yule", *Journal of the History of Biology* 37:73-101.

Tabery, James. 2006. "Looking Back on Lancelot's Laughter", *Mendel Newsletter* 15: 10-17.

Tabery, James. 2007. *From a Genetic Predisposition to an Interactive Predisposition: Reconceptualizing and Reframing the Ethical and Legal Implications of Gene-Environment Interaction Research on Antisocial Personality Disorder*. Master's Thesis. Pittsburgh: University of Pittsburgh.

Tabery, James. Forthcoming. "$G \times E_B$ and $G \times E_D$: Looking Back, Moving Forward", *Development and Psychopathology*.

Taylor, N. B. 1931. "The Relation of Temperature to the Heart Rate of the South African Frog (*Xenopus Dactylethra*)", *Journal of Physiology* 71:156-168.

Thompson, E.A. 1990. "R.A. Fisher's Contributions to Genetical Statistics", *Biometrics* 46:905-914.

Tinbergen, Nikolaas. 1963. "On the Aims and Methods of Ethology", *Journal of Comparative Ethology* 20: 410-433.

Toulmin, Stephen. 1977. *Human Understanding: The Collective Use and Evolution of Concepts*. Princeton: Princeton University Press.

Vreeke, G. J. 2000. "Nature, Nurture and the Future of the Analysis of Variance", *Human Development* 43: 32-45.

Waddington, Conrad H. 1957. *The Strategy of the Genes*. London: Allen and Unwin.

Wagman, Paul. 1973. "The Brains Do Battle in I.Q. Controversy", Pp. 18, 28 in *The Boston Phoenix*, November 6, Boston, Massachusetts.

Wahlsten, D. 1990. "Insensitivity of the Analysis of Variance to Heredity-Environment Interaction", *Behvioral and Brain Sciences* 13: 109-161.

Wahlsten, D. 1994. "The Intelligence of Heritability", *Canadian Psychology* 35: 244-260.

Wahlsten, D. 2000. "Analysis of Variance in the Service of Interactionism", *Human Development* 43: 46-50.

Waters, C. Kenneth. Forthcoming. "Causes that Make a Difference", *The Journal of Philosophy*.

Wells, G. P. 1978. "Lancelot Thomas Hogben", *Biographical Memoirs of Fellows of the Royal Society of London* 24:183-221.

Wells, G. P. n.d. "Father and Son", Lancelot Hogben Papers (A.38), University of Birmingham.

Werskey, Gary. 1978. *The Visible College: The Collective Biography of British Scientific Socialists of the 1930s*. New York: Holt, Rinehart and Winston.

Westfall, Richard. 1971. *The Construction of Modern Science*. New York: John Wiley and Sons, Inc.

Williams, Roger John. 1969. *Biochemical Individuality*. Austin: University of Texas Press.

Wilson, Margaret. 1999. *Ideas and Mechanisms*. Princeton: Princeton University Press.

Wimsatt, William C. 1972. "Complexity and Organization", in K. F. Schaffner and R. S. Cohen (Eds.), *Proceedings of the Philosophy of Science Association*. Dordrecth: Reidel, pp. 67-86.

Wimsatt, William C. 1976. "Reduction, Levels of Organization and the Mind Body Problem", in G. Globus (Ed.), *Consciousness and the Brain*. New York: Plenum Press, pp. 205-267.

Wimsatt, William C. 1984. "Reductive Explanation: A Functional Account", in E. Sober (Ed.), *Conceptual Issues in Evolutionary Biology*. Cambridge, Mass: MIT Press, pp. 369-385.

Winton, Frank. 1927. "A Contrast between the Actions of Red and White Squills", *Journ. Pharm. Exper. Ther.* 31:137-144.

Woodward, James. 2002. “What is a Mechanism? A Counterfactual Account”, *Philosophy of Science* 69: S336-S378.

Woodward, James. 2003. *Making Things Happen: A Theory of Causal Explanation*. Oxford, Oxford University Press.

Yates, F. 1964. “Sir Ronal Fisher and the Design of Experiments”, *Biometrics* 20:307-321.

Yates, F., and K. Mather. 1963. “Ronald Aylmer Fisher”, *Biographical Memoirs of Fellows of the Royal Society of London* 9:91-120.

Yule, George Udny. 1902. “Mendel’s Laws and Their Probable Relations to Intra-Racial Heredity”, *The New Phytologist* 1:193-207, 222-238.